Vol. 63, No. 1 January, 1956

Psychological Review

THEODORE M. NEWCOMB, Editor
University of Michigan
Lorraine Bouthilet, Managing Editor

CONTENTS

Affect and Perceptual Learning.....Gardner Murphy 1

An Axiomatic and Experimental Study of Sensory Order and Measure.....Eugene H. Galanter 16

"Functionalism" in Perception.....W. C. H. Prentice 29

Extinction Re-examined and Re-analyzed: A New Theory.....Gregory Razran 39

Issues Distinguishing Idiographic from Nomothetic Approaches to Personality Theory.....John L. Falk 53

A Micromolar Approach to Behavior Theory.....Frank A. Logan 63

Subception: Fact or Artifact?.....Charles W. Eriksen 74

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

CONSULTING EDITORS

SOLOMON E. ASCH
ROBERT R. BLAKE
STUART W. COOK
CLYDE H. COOMBS
LEON FESTINGER
W. R. GARNER
JAMES J. GIBSON
D. O. HEBB
HARRY HELSON
E. R. HILGARD
CARL I. HOVLAND
E. LOWELL KELLY
DAVID KRECH
ROBERT W. LEEPER
KENNETH MACCORQUODALE
ROBERT B. MACLEOD
DAVID C. MCCLELLAND
GEORGE A. MILLER
GARDNER MURPHY
OSCAR OESER
CARL PFAFFMANN
CARROLL C. PRATT
JOHN P. SEWARD
DAVID SHAKOW
RICHARD L. SOLOMON
ELIOT STELLAR
S. S. STEVENS
ERIC TRIST
EDWARD L. WALKER
ROBERT W. WHITE

The *Psychological Review* is devoted to theoretical articles of significance to any area of psychology. Except for occasional articles solicited by the Editor, manuscripts exceeding twelve printed pages (about 7,500 words) are not accepted. Ordinarily manuscripts which consist primarily of original reports of research should be submitted to other journals.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Tables, footnotes, and references should appear on separate pages; all of these, as well as the text, should be typed double-spaced throughout, in all manuscripts submitted. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail. Manuscripts should be addressed to the Editor, Dr. Theodore M. Newcomb, Doctoral Program in Social Psychology, University of Michigan, Ann Arbor, Michigan.

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STREETS, LANCASTER, PA.
1333 SIXTEENTH ST. N. W., WASHINGTON 6, D. C.

$8.00 volume $1.50 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized Jan. 8, 1948

Send all communications, including address changes, to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must arrive by the 15th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright © 1956 by the American Psychological Association, Inc.

VOL. 63, NO. 1 JANUARY, 1956

THE PSYCHOLOGICAL REVIEW

AFFECT AND PERCEPTUAL LEARNING

GARDNER MURPHY
The Menninger Foundation

A question directed to a group of modern psychologists regarding the role of affect (or the physiological substrate of affect) in perception might elicit the following responses: (*a*) Affect cannot influence perception. (*b*) Affect can intensify or weaken the perceptual response (e.g., through a mechanism like that of dynamogenesis or intersensory processes). It may intensify, or prolong, or consolidate the percept. It may prevent the perceptual structure from changing (this is sometimes called rigidity), or it may make it more labile, but it may not directly alter perceptual structure. (*c*) Affect can intensify or weaken certain aspects of a perceptual response, as in some of the "perceptual defense" studies. (*d*) Affect can intensify or weaken certain aspects of a perceptual response in such a way as to influence later perceptual responses, e.g., it can result in perceptual learning. The present paper explores some of the issues related to the last of these possibilities.

I

It is interesting to note the unfinished debate (cf. Socrates vs. the Sophists, and the Gibsons [4] vs. Postman [15]) between those who assert and those who deny that the development of perception must move toward realistic contact with the environment. But in view of the fact that in everyday life, and in our laboratories, our hopes and fears lead to appearances that are notoriously deceptive, it seems to the present writer a little odd that affect does not enter into these discussions of perceptual development (or rather, I should say that affect is not officially included among the factors influencing the perceptual act).

This becomes still more strange when one notes that "set," "hypothesis," "attitude," "mood," and "value" are almost universally agreed upon as influencing perception. When one focuses on these processes so as to see them at close range, they are seldom affect-free. And the influences which they exert are, so far as I know, universally conceded to last on into subsequent perceptual responses, as has long been clear, especially since the time of Rubin (23).

The issue here is not whether perception moves toward greater *realism* or toward *better organized* structure, or toward finer differentiation and fuller integration. Space precludes discussion of these and many other possibilities. The present paper presupposes that the types of perceptual learning described by the Gibsons exist. Rather, the aim is solely to explore the rather neglected but very pressing issue regarding the role of affect in perception and in perceptual learning. The only aspect of affectivity considered here is the pleasantness-unpleasantness dimension.

II

In exploring the possibility that affect may alter the emphasis, or dominance relations, within a perceptual act, the following seem to be some of the more easily identifiable processes: (*a*) the percept may be so structured as to reveal emphasis upon a component which has been associated with positive hedonic tone (autistic perception, i.e., a drive-satisfying type); (*b*) the percept may be so structured as to cope with a difficult or challenging environment by emphasizing a component associated with negative tone, or, in any event, with components acting as a threat, requiring an "alerting" response. It will be noted that "alerting" likewise operates in the service of a "pleasure principle" by warding off that which might cause distress. It constitutes a roundabout way to pleasure (as in "secondary process thinking" in psychoanalysis); (*c*) a percept may be so structured as to be more satisfying, e.g., more interesting in its own right, not because it represents the environmental possibilities to be more favorable than they actually are, but through the greater satisfaction that comes from grasping more fully, rather than less fully, the nature of the stimulus. Here belongs, for example, the aesthetic satisfaction of hearing a full orchestration rather than a few tuneful bits of symphonic music. Perhaps we should add, here, the satisfaction of being independent, not being suggestible, etc., but the case for satisfaction in making contact with reality remains. (*d*) A percept may become rigid through a circular-response mechanism (remotely reminiscent of causalgia, or fingers clutched about a live wire). The relation of such rigidity to repetition-compulsion may be worth exploring. In many instances, however, the rigidity is not a simple reflex but expresses a fear of giving up the percept, lest we be lost. This is especially evident when one cannot give up a rigid self image, through fear of disorientation or having nothing left to cling to. The clinically minded reader will think of many other categories. (*e*) A percept may be consolidated in such a way as to hurt the perceiver, through the mechanism described by Freud as masochism; the punishment gives satisfaction.

Most (if not all) of these mechanisms, whatever their ultimate form, are most easily conceptualized in hedonistic terms; they all seek pleasure or release from stress. The emphasis of the present paper is on pursuing just *one* of the present possibilities, namely, the possibility wherein perceptual learning is of a simple, directly pleasure-seeking type, in which the perceptual structure is so organized as to accent or emphasize those components which directly give pleasure. At a later time I hope to be able to do more justice to those instances in which the emphasis is carried so far that something gets *left out*, and there is consequently distortion, evasion, misinterpretation.

In earlier studies some of us (8, 25) have used Rubin's terminology of "figure and ground" to describe the emphasis or dominance factor within a perceptual whole. This use may or may not be advisable. In the present paper the concept of "emphasis," as used by Tolman (30) and Muenzinger (13) will be used, and an attempt made to suggest that we are basically dealing with the same thing as "dominance," as used by Razran (18), or "regnancy" as used by Murray (14). That particular form of dominance known as "figure and ground" will be noted only when the situation resembles that of Rubin.

I suggest two hypotheses: (*a*) subject to exceptions to be noted later, perceptual structure changes in such fashion as to give *emphasis* to that which is

pleasant, in contrast to that which is unpleasant or affectively neutral; (*b*) many such changes are *cumulative;* constituting a form of perceptual learning. This means that directional modifications in cognitive structure are made as a result of affective factors. This would bring us rather closely into line with Chein's (10) definition of autism as "the movement of cognitive processes in the direction of need satisfaction." It is suggested that positive affect not only plays a vital role in the *modification* of perceptual responses, but that it is a clue to some forms of perceptual learning.

Postponing for the present the detailed discussion of negatively toned components in the perceptual field, the hypotheses mean that, given two or more aspects of a perceptual whole and repeatedly associating one of these with positive affect, will tend, when these various perceptual aspects are simultaneously presented, to lead the dominant role to be taken by that which has been associated with positive affect.

We assume, of course, that our learner, if awake, is actively forming these associations; shifting his attention; interested in this detail, bored with that one; and making such sense as he can out of his situation. Perception is an integrative process or act, rather than an unselective intake of stimuli. What we are concerned with here is rather a specific empirical relation between the various aspects of the perceptual whole, namely, the apparent fact that positive affect (pleasantness) may serve to modify a perceptual response by accentuating some aspects of it.

III

While my purpose is not chiefly historical, I should like to give a few illustrations of recent efforts to cope with this problem. Woodworth (32) has pointed out that we should regularly recognize an affective factor in learning and should describe this as "reënforcement of perception," rather than reinforcement of an *act*. In the course of experiments in salivary conditioning, Razran (20) showed that paintings which had one affective status for a given hungry subject took on a different aspect (apparently because of a new accent on some features and the loss of accent on other features) as a result of a food stimulus paired with them. Sanford (24) demonstrated that the meaning of photographs as indicating objects related to food was altered by periods of food deprivation. Schafer and Murphy (25) devised a simple situation involving two caricatured faces, one face associated with a reward, the other with a punishment, and showed that when the two faces were presented together, the one that had been previously "rewarded" took on the dominant role. This experiment, when repeated in a modified form by Rock and Fleck (22), gave nonconfirmatory results; but Jackson (8) has repeated both the Schafer-Murphy and the Rock-Fleck experiments and showed that among other things the field within which such modifications are to be induced must be small and compact, the Schafer effect being easy to reproduce when the Schafer technique is used. Incidentally, one of the most interesting things in the Jackson study was the evidence for intraindividual consistency. Those subjects who showed this effect more plainly with one pair of faces tended to show it likewise with another pair of faces. Smith and Hochberg (28) have shown that the Schafer effect can be reproduced without using reward; they utilized electric shock as punishment and found that the shock had a tendency to prevent the shocked aspect from appearing as figure. This raises the old question of whether the absence of punishment is in itself a "reward."

Yet in the huge welter of experiments and communications appearing in the "New Look" psychology and the complex issue of the role of affect and impulse upon perceptual and other cognitive structures, I am not aware that much is known about the specific way in which the perceptual field is altered by affect. It is true that much has been written about perceptual defense. But in perceptual defense we have a process of exclusion, and the actual dominance relations within the perceptual field are not investigated.

IV

I should like first to stress the importance of *instability* in perceptual structure, if modification is to occur. In attempting a preliminary answer to this question, James Simpson, in helping me with this paper, has suggested a diagrammatic way (Table 1) in which to show a hypothetical course of perceptual development under the influence of various needs and through the utilization of rewards (stimuli bearing positive affect). In Table 1 we are to imagine our hypothetical subject as presented tachistoscopically with a complex visual schema including the numbers 1–4–9 embedded in other material. The first response is one of curiosity, which in the light of the work of Harlow and his associates (5) I should like to call both a cognitive and an affective response, and a social response to the experimenter. The experienced affect is perhaps at first a slight anxiety with mixed feeling tone. The perceptual response is the number sequence 1–4–9, an undistorted percept. The subject at this time makes no overt response because the initial percept is unstable and fleeting. We proceed to a second presentation. There is a need to achieve, as the subject warms up to the task, and also a persisting need to solve the problem

TABLE 1

HYPOTHETICAL COURSE OF PERCEPTUAL DEVELOPMENT UNDER THE INFLUENCE OF VARIOUS NEEDS AND THROUGH UTILIZATION OF REWARDS (STIMULI BEARING POSITIVE AFFECT)

Ambiguous Tachistoscopic Stimulus	Sequence in Trial 1	Need Pattern (conscious and unconscious)	Predominant Affect (kind/amount)	Perceptual Responses	Overt Responses
X1X4X9X	1st	Cognitive: (curiosity) Social: (experimenter)	Anxiety, Unpleasant and Pleasant/Minimal	149 (undistorted percept)	None
	2nd	Achievement and Cognitive (solve problem) Social: (experimenter)	Pleasant and Unpleasant/Moderate	1492 (cognitive distortion)	None
	3rd	Sex and Social (girl friend)	Erotic: Pleasant/Strong	5-1492 (affective distortion)	None
	4th	Achievement and Cognitive (solve problem) Social: (experimenter)	Pleasant/Moderate	149 (cognitive correction)	"149" (correct solution)
					(Subject receives $ reward)
	In Subsequent Trials:				
X1X4X9X	1st	Acquisition (money) Achievement (solution) Social: (experimenter)	Pleasant/Moderate	149 (learned interpretation)	"149" (correct solution)

and to make a social response to the experimenter. This mixture of pleasant and unpleasant components will be a factor in unsteadying the percept, and there are variations around the 1–4–9 response. For example, in this instance the instability leads to the addition of an additional number 2, since 1–4–9–2 is one of the overlearned stock number combinations available near to consciousness and to the effector response system. On the third presentation, however, a slight shift of internal conditions causes our subject to remember his girl friend with considerable delight in thinking of a coming date, and her telephone number 5–1492 is perceived. Still, however, there is no overt response. On the fourth presentation, the cognitive and achievement and social factors remain. There is a pleasant lingering affect. The reality trend has swung again slightly into the ascendant and the cognitive response 1–4–9 is now called for and uttered, and the subject receives a reward. In subsequent trials the first thing to be noted is the re-enforcement effect of the reward and the further structuring of the response 1–4–9. Similar sequences can be devised for other patterns.

Our interest is in the question: where and how does the reward get in its work? The issue which seems to me to call for attention is the fact that initially there are *unstable* perceptual responses which permit modifications—sometimes only a shift in dominance relations, sometimes so gross and heavy an accent on one detail that other aspects of reality are missed, and there is distortion or error. It is by virtue of the momentary instability of the perceptual system that the reward can exert its influence. Doubtless my own autistic distortions are responsible for my assigning importance to these issues; but I should like to raise a series of questions upon which clarification may perhaps be achieved through the critique, controversy, and experimentation that may be stimulated.

V

Suppose we ask ourselves now just *how* can an unstable perceptual response system be *modified?* It may be modified either through changes in the world of outer stimulation or through changes inside the organism. Under the former head, absolute or relative increases or decreases in the intensity of the stimuli may alter the perceptual field. Some aspects of the field may be muted, or may even be eliminated. Other aspects, therefore, may be relatively or indeed absolutely emphasized. This is what we mean by an alteration in the relative emphasis in various parts of the field. Secondly, it is my thesis that stimulation from within the organism (e.g., from various internal sources, including neural and endocrine sources) may at any time accent or mute some aspects of the perceptual field. It is suggested that affective processes, whatever their physiology, or however well we may learn to gauge their presence by the galvanometer or other instruments, should regularly exert some influence. The influence might be of a purely intersensory type, as a color may change appearance during the action of a sound or tactual stimulation; but at any rate, it is hard to see how affective processes could be so isolated from the functions of the organism as to exert no influence at all on the perceptual response system. Granting the great importance of the structural properties of the environmental stimulus and the structural properties of the responding organism, affectivity nevertheless appears to involve a continuous input of stimulation, resulting in facilitation or inhibition. A commonplace illustration is the fact that we are so delighted with something that we do not notice an imperfect de-

tail, or are so angry that we "look for trouble." We must therefore ask more precisely about the processes that might lead affectively toned material to take a dominant place in perceptual structure.

Let us note some relations of affectivity to body movement (the following are commonplace evolutionary assumptions; cf. for example, Holt, 6). Movement toward and movement away from objects are present even in unicellular animals. As organization becomes more complex, the means of locomotion and of contact and avoidance become more complex, but the basic differentiation between "movement toward" and "movement away from" remains. Let us assume that primitive awareness, when once it begins, involves some capacity to differentiate stimulus objects from their contexts. Moreover, insofar as we note that primitive behavior is oriented positively toward some elements in the environment and negatively toward others, let us assume that there are primitive impulsions toward and away from objects. From the present viewpoint, these primitive impulsions toward and away from objects may involve affects. Positive hedonic tone may appear at the level of experience when we are approaching an object and getting more and more contact with it. Some objects are biologically beneficent. *In the long run, from an evolutionary point of view, we should expect positive movement to occur in the case of those objects which it is "good for" the organism to possess. By the same logic, there is a primitive basis for positive affect; positive affect is the subjective aspect of approach behavior.*

The types of biologically determined perceptual responses that we have been assuming are of course enormously elaborated by the learning process, and in particular by the education of human beings in the world of symbols. Objects that are at first devoid of great affect (words, gestures, coins, etc.) come to carry enormous affective freight, and one's seeing and hearing may be influenced by the *meaning* of the symbol, i.e., the symbol may take on a dominant role because of what it stands for. But it is not clear that the basic functioning of *affect* could be altered by this fact. Indeed, this fact seems to point to our human dependence on abstract or remote objects by virtue of their linkage, in our life history, with more concrete and immediate goals. The dominance relations within a perceptual field may indeed be determined largely by very remote, subtle, and complicated acquired meanings of the components of the field. But the hypothesis is offered that however the affective character of these components of the field arises, their affective character influences their dominance relations.

When, therefore, at the level of direct satisfactions of biological needs, objects begin to stimulate the individual, there is a tendency for the perceptual response to the satisfying object to become more and more prominent. Primitive movements of approach are therefore accompanied by primitive positive hedonic tone. Perceptual responses to objects that are pleasant tend to move into prominent positions in the perceptual structure.

It is agreed that expectancy and hypothesis testing are fundamental in perceptual learning. But why does expectancy in relation to the pleasant become dominant at times over expectancy in relation to the neutral or the unpleasant? Perhaps part of the problem is the means-end relationship; i.e., we look for trouble if this helps us to *avoid* trouble. But if we are simply functioning at a phantasy level, or if, in a post-test series, we are simply ruminating on the good, bad, and indifferent things through which we have passed, there is

no means-end service accomplished by expectancy related to the neutral or the unpleasant, and the pleasure principle can operate.

From an organism's general evolutionary position, response to the "good" and to the "dominant" in the environment are often identical. We must not miss the good things that are available. Expectancy leads to sensitization. But carry-over of this good-expecting attitude can be very disruptive of adjustment; when these attitudes carry over into situations in which life is rigged against us, we get into trouble with reality, since perception has to be basically a reality-oriented process. It is for this very reason that perceptual responses become rather deeply ingrained, and for this very reason that there is a danger that perception will miscarry at times and lead us to expect the good even where it is not there. We know at heart, as Ulysses did, that we have to be lashed to the mast, or the sirens and the lotus eaters will get us.

From this standpoint it is believed logical to conclude that during continuous or repeated stimulation, pleasant things tend to receive more and more emphasis in any unstable perceptual field. I cannot here discuss the question whether such hedonic factors in learning operate in essentially the same way at this primitive level and also at the level of such complex phenomena as self-defense, self-justification, and the intrinsically satisfying process of rational planning. Razran (18, 19), among others, has amassed much evidence that there are levels of learning. I have, however, indicated (page 2) where my bias lies, namely, in the belief that a pleasure principle enters into much that at first sight seems free of it; it seems to me that the perceptual field is structured by the satisfactions of the social and cognitive worlds, e.g., through the satisfactions which thinking subjects necessarily confront through the joys and the vicissitudes of the thought process itself. I cannot, therefore, find an area where hedonistic perceptual theory cannot apply.[1]

Positive affect will be present whenever we are carrying forward evolutionarily given goal-seeking activities. The satisfactions, of course, may be aroused either directly by the goal or by objects associated or conditioned to the goal, and the process may be quite complex, as in being gratified through money or honorific letters after one's name. The hypothesis entertained is that any aspect of a perceptual field which is signalized in this fashion by gratification is likely to be the conscious aspect of a goal-seeking activity. Moreover, in the light of the general theory of dominance in the theory of conditioning, response to that which is *most* gratifying will tend to have the highest position in the dominance hierarchy. The greater the satisfaction, therefore, the greater the movement, during successive encounters with a situation, toward the apex of the dominance hierarchy.

If we now ask by what process a component in a perceptual whole can be reinforced by the sheer presence of affect, we may tentatively restate the problem as one ultimately involving the dominance relations which may perhaps hold, not only in conditioning, but in all psychological events, as shown in Fig. 1. From Razran's extensive analysis (18) it appears that conditioning does not simultaneously attach S_1 to R_2 and S_2 to R_1; the dominance relation is expressed in the fact (ordinarily, at least)

[1] It would be of interest to see whether the language of *drive reduction* (and of anxiety reduction) could state all these things better. I am inhibited in attempting this myself, as I often find it difficult, in my own case and in that of others, to be sure about drives (and anxieties), while pleasantness and unpleasantness usually remain easily identified.

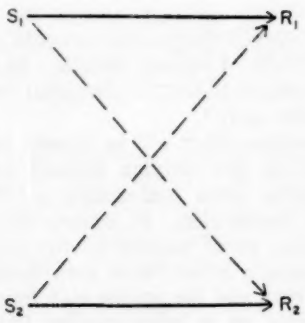

FIG. 1. Dominance relations.

that one of these possibilities is realized, not both.

Let us at the same time use the Birch-Bitterman (2) concept of "sensory integration." Sensory integration (and disintegration or loss of connection) goes on continuously. If the question be raised, "How can such integration be achieved by an affective component?", our reply is that the affective component often betrays the dominance relations at work in the response system, which as we have seen is both an *affective* and an *overt* behavior system. For example, when you say, "Tell me what you see," the subjects in the experiments cited above moved in the direction of emphasizing that aspect of the field which had previously been associated with reward. Sensory integration takes one form through the action of affect and simultaneously excludes other forms. The results of such association should be cumulative, as are the effects of most types of learning involving simple association; and the experimental data, as in the Schafer-Murphy and Jackson experiments, confirm this view.

VI

Under the very simple conditions described so far, it is sufficient to say that that which is not satisfying is that which is shoved to one side. Perhaps the biological process of escaping or shoving to one side is the basis for the phenomena of negative affect. The turning away or rejection of an experience at the conscious level corresponds, from the present viewpoint, to the biological turning away or rejection of the object. This would apply even to the Smith and Hochberg experiment in which electric shock was sufficient basis for the associated face's being rejected. I suggest that the perceptual field takes on a structure in which the dominant role is played by that aspect which is being pursued, and ground by that which is being fled.

But, as everyone knows, there are important instances of dominant roles played *not by the pleasant, but by the unpleasant*. Charles Snyder has phrased this simply and well: "How is an organism able to move *away* from a given object without that object's first becoming dominant in the perceptual field?" A problem that cannot further be avoided! Obviously, the whole of the present thesis, if correct at all, can apply only to a *trend during successive interactions with the environment,* not with a process full blown when it first appears. It is only with experience, if ever, that structure becomes stabilized in a satisfying way. Perception is often unstable, wobbly, during the learning. Moreover, the unpleasant is at times dominant *both at the time of its first appearance and after the achievement of stable structure.*

I have indeed taken on a rather large problem when stating that that which is associated with unpleasantness is among the things tending to become nondominant. In some situations that which is most frustrating is actually most likely to be observed. This is analogous, of course, as already noted, to the "emphasizing" (or alerting) effect. It is often touch and go whether we shall *face* or *avoid* a threat stimulus. Since

in this paper I am not allowing myself to dilate upon all the conditions (expectancy, masochism, etc.) that might lead to such a result, I will oversimplify by referring to the asymmetrical **U** curves which have sometimes resulted from studies of "recall of pleasant and unpleasant experiences" (**11**, **12**, cf. also **16**). These curves, ideally taking perhaps such a form as is shown in Fig. 2, may represent simultaneously the advantage of pleasant over unpleasant, and the advantage of strong over weak affect. *Severely unpleasant* alerting situations can take precedence over *mildly pleasant* situations. That which becomes dominant may be that which is very annoying, rather than that which is mildly pleasant. Indeed, I gladly agree that there are many instances analogous to the phenomenon of trauma, in which the alerting affect may have consequences far *more* marked than those of positive affect. Various types of **U** curves would therefore be appropriate to depict the situation. This means that we must include a dimension of sheer *intensity* along with our affectivity dimension. And, to be logical, we must again say that the unpleasant gets into the picture in proportion to its intensity, and because it is ultimately a means to the pleasant.

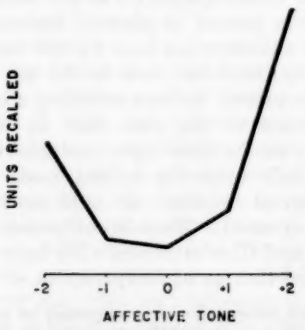

FIG. 2. Advantage, of pleasant over unpleasant, and strong over weak affect.

In the recent studies by Dulany (**3**) and Pustell (**17**) another dimension of the punishment situation is stressed. Dulany appears to have shown that when one of four tachistoscopically presented figures is severely shocked, and *S* learns to avoid *reporting* that this one "stands out" among the rest, he also learns to avoid *perceiving* that this one "stands out"; the result appears to be experiential as well as behavioral. But in Pustell's parallel study in which the shock comes so quickly after the stimulus that *S cannot avoid it*, many *S*s (notably the males) develop "vigilance" (lowering of thresholds for the critical stimulus) rather than defense. Such studies may help us to understand the omnipresent importance of *set* in perceptual learning. Set can either lower a perceptual threshold or raise it. The question is why we are sometimes set for the pleasant, or, as in the vigilance experiments, why we are set for the unpleasant. Maybe because we can thus better *use* the pleasant or thus *get away* from the unpleasant. From such a viewpoint, needs and readinesses are the same thing. The conflict between hypothesis and need psychology disappears.

The discussion of the conditions governing the dominance of the unpleasant could lead into all outdoors, and I must call a halt. But let me emphasize once again that the present approach recognizes that the unpleasant often can, and does, become dominant.

VII

It is time now to look at those issues, among the ones presented, that can be decided by experiment. The following questions must be answered in the course of any investigation dealing with these issues:

Under what conditions does the "Schafer effect" (perceptual learning

process resulting from dominance of the pleasant) appear?

1. In which sensory modalities?
2. Does it appear when reward alone is used, or only when reward and punishment are combined?
3. At what levels of awareness does it appear?

a. For example, does it appear when affectively toned stimuli are presented below their usual (sensory) threshold value?

b. Does it appear when affective material is supraliminal, but no explicit connection is formed by *S* between affect and the rest of the perceptual field?

c. Does it appear *only* when an explicit connection is formed?

d. Does it appear *only* when the connection is verbalized?

e. Does it appear *only* when there is ego involvement in forming the connection?

4. What is the role of frequency, intensity, vividness in such learning?
5. What is the role of extinction (and of consolidation) when rewards and punishments are removed?
6. Does this kind of approach have any utility at all when the satisfactions come chiefly from events inside the person—as related, for example, to self-respect? How is the reappraisal of oneself during psychotherapy related to the simplest figure-ground dynamics?
7. Last, but by no means least, what is the relation of *extrinsic* rewards and punishments to the *intrinsic* satisfactions which come directly from clear, internally consistent, well-oriented perception, and the frustrations which come from confused, bewildered, internally inconsistent perception? If there is such a thing as joy in discovery, there is such a thing as joy in the sheer use of the perceptual apparatus, and this must be carefully distinguished from the delights of autistic distortion, if for no other reason than that at a later time we shall have to compare the satisfaction of reality seeking with the satisfaction of reality evasion.

The experiments about to be described comprise a group of interlocking studies in the visual, tactual, and auditory modalities, utilizing adult subjects who can, when desirable, take part in all of the various experiments and in other experiments and clinical studies at the Menninger Foundation.[2]

I will sketch here a few of the studies which my associates and I are carrying out in the attempt to test such hypotheses as these and to develop methods suited to the definition of affect and its role in perceptual learning.

After Jackson's repetition of the Schafer-Murphy experiment, one of the more obviously pressing needs was to ascertain whether similar results could be achieved in other modalities. After a large number of pilot studies introducing variations in the basic design of the Schafer study, we finally decided to stick as literally as possible to the bare outline of the Schafer method, but to use the auditory and tactual modalities. In the auditory field the technique finally developed was to tape record two voices, each reading prose material (from Rachel Carson's *The Sea Around Us*); to present in planned haphazard order the materials read by one voice, and by the other; then in the test series to present by tape recording a conglomerate of the two—that is, both voices on the same tape—and ascertain to which voice the subject made his perceptual response (he could not hear both at once). These investigations by Fred and Charles Snyder (29) have indicated that an auditory version of the

[2] This investigation was supported by a research grant, M-715(C), from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

Schafer effect is feasible: the subjects tend to hear those materials which were associated with monetary rewards in competition with those associated with monetary punishments.

In the investigation of tactual figure-ground effects by Teodoro Ayllon and Robert Sommer (1), ceramic plaques were used, with a deep irregular contour line vertically down the middle. In the training series, the subject traced the groove with his gloved right index finger, while looking at a right-pointing face painted to the left of the groove or a left-pointing face painted to the right of the groove, and simply learned the names of these faces while receiving electric shock for one face and no shock for the other. (The absence of shock appears to be a reward, but this controversial question will have to be studied much more fully at a later time.) In the *test* series, the subject, now blindfolded, moved his gloved finger down the groove, which (except for set-breaking figures) was always the same. Here there was a trend to perceive the nonshocked face (a "defense" reaction in accordance with the autism principle) when the unpleasantness of the shock was rated by *S* as "slight."

When, however, the shock was at a level reported to be "moderate" or "severe," there was a significant tendency toward "alerting," i.e., to perceive the *shocked* face. Further studies with both auditory and tactual figure-ground effects are being carried forward, with emphasis upon three problems: first, the problem of *generalization*, when new material is used that has something in common with the material on which training has already been carried out; second, the "extinction" of freshly acquired perceptual habits; and third the defense and "alerting" responses at different levels of unpleasantness. Several other half-finished studies of perceptual dominance are under way, designed to answer some of the questions stated on page 10.

We have also been interested in the use of the Ames techniques (9), especially with reference to the question: Can the perceptual distortions or misinterpretations investigated by the Ames methods be reduced or eliminated to some degree by the use of affective factors related to certain cues; can the unlearning of such misinterpretations be facilitated by a process through which distortion is punished and correct perception apprehended? James Simpson (26) selected the aniseikonic-lens situation as one serviceable for such studies. He has apparently established the fact that positive and negative affect, induced, respectively, by music and a ratchet sound, operate to accentuate or weaken the distortion effects of the very complex visual field which the aniseikonic lenses present, so that acceptance or rejection of the usual aniseikonic effects is differentially facilitated. Insofar as the test situation, after music and ratchet sound have been discontinued, presented the subject with a complex and conflicting set of cues for space perception, we may say, by a somewhat forced use of language, that this was also a study of the influence of affect upon figure-ground organization. It is probably better, however, to avoid such linguistic forcing, and simply to say that the interpretation of the cue structure depended upon affects which had been aroused in the earlier (learning) situation.

Perhaps it is worth while to say still once again that we do not believe that "autism" effects are universally demonstrable. We do not believe that affect is the only factor in perceptual learning. We do not believe that affect operates in the same way in all persons. We do not believe that age, sex, and cultural factors can be ignored. Indeed, the structure of the affective life,

person for person, has proved to be a major problem to cope with in such studies. We find, for example, that certain effects are relatively easily obtained with naive subjects, but in general very uncharacteristic of subjects who understand such experiments and introduce problems and hypotheses of their own regarding the processes at work. As would be suspected on the basis of the material reported by Razran (20), the perceptual changes belong at a certain level of learning, and characterize particular types of orientation to a task. Our question is not: "Why do our colleagues resist extensions of the evidence for autistic phenomena?" but simply: "*Under what conditions* does the Schafer effect appear?"

VIII

Indeed, working as we do in a clinical atmosphere, side by side with those whose concern is with the endless richness of differences in human needs and cognitive patterns, and working with the "perception team" of George Klein, Riley Gardner, Philip Holzman, and Herbert Schlesinger, we have been constantly alerted to the problem of individuality. While our chief concern is with the general laws stated above, this does not mean that individual differences can be ignored. We have no choice, for facts about individuality leap at us and upon us. We shall attempt, then, to show that the reality of individual differences does not invalidate our belief that general laws are our first problem.

The fact seems to be that among subjects in most such experiments—young adults of little more than average sophistication—there are always a few who appear consistently to move "in the unexpected direction." Long ago, Razran (19) pointed out that some of his subjects in simple adult human conditioning situations moved in the expected direction "like Pavlov's dogs," that some did not condition, while others conditioned "in the unexpected direction." Others of the New Look investigators have, of course, been finding this sort of thing. We are inclined to think that the following problems take shape in consequence of this fact: (*a*) Some of the "rewards and punishments" as seen from the experimenter's point of view are simply not rewards and punishments for the particular subjects. To expect such "rewards and punishments" to have exceptionless consequences would be like expecting a color-blind man to respond in a normal way to a rainbow. In Jackson's study, those *Ss* who admitted feelings about the money rewards and punishments showed a very marked tendency to move in the rewarded direction, while those who denied feelings about the money showed a tendency to move in the opposite direction. (*b*) Some of the rewards and punishments are not at that point on the **U** curve to which we assigned them. They are either intensely gratifying, bring in ego involvements—"see how smart I am," etc.—or they serve as threats and challenges, as alerting stimuli which may bring negative affect when we expect positive, etc. "What is he trying to do?" "Is he one of these psychologists that always has a trap for you?" Or, most commonly, "What is the point of this shoving money at me and taking it away?" All these considerations mean that even money, which is the most ancient and general of human inventions designed to provide something that has a dependable affective value throughout a community and can be standardized on a hedonistic basis as something that one strives to win and hates to lose, does not induce affect of dependable uniformity. If we can find rewards and punishments that will better serve the purpose, we shall be glad to use them;

indeed, we have interesting results using unpleasant tones, electric shock, mild frustrations in carrying out intellectual tasks, and other annoying stimuli, as well as music stimuli, empirically "calibrated" for affect.

In general, however, we are inclined to believe that it does not make a great deal of difference what positive and negative stimulation we use, provided that the affect is fairly uniform. Under these conditions we appear to be dealing not with genuine exceptions with reference to affectivity in connection with perceptual dominance, but rather with the fact that sometimes no affectivity is present, and sometimes the affectivity which we assumed to be present was replaced in some subjects by a different kind of affect. If such affective categories are empirically determined, they must be freshly tried out through "cross validation."

Cross validation of this sort is feasible. It is at present our impression that those subjects who move in the direction of punishment in some experiments tend to do so in other such experiments to a degree greater than would be expected by chance alone; we seem to encounter the recognizable "negative" personality. The data seem, however, to indicate that the exceptions are cases where quantitative and qualitative differences between subjects result in different affects being invoked and necessarily different consequences involved. A needed experiment from the point of view of sound, clear generalization is one in which the subject has no chance to develop these negative responses because he never clearly apprehends the components in the situation that are being rewarded and punished, and therefore cannot effectively defend himself against them. Some of our experiments seem to offer the possibility of achieving this purpose. But we agree that our experiments are *very far indeed* from offering adequate explanation of individual differences, or the reasons for them.

Another related issue: To what degree are our results attributable to the training of our subject in certain habits of attending? The question is related to the problem of whether results of these sorts can be due to distraction. Neither in the training series nor in the test series are the data adequate regarding the deployment of the subject's attention. We are exploring the question of whether the whole Schafer effect can be conceptualized as the "conditioning of attention." Another urgently needed branch of experimentation, in which we are following the suggestions of J. E. Hochberg, is one in which the subject's perceptual learning reaches a point at which he cannot, despite great effort, perceive what has been punished, or cannot avoid seeing what has been rewarded.

We hope of course, that such experiments will enable us to understand better the nature of perceptual dynamics, the nature of the learning process, the nature of autism, and some of the possibilities for educational and clinical application.

It looks to us at this time as if Woodworth were right and as if we were reinforcing perception. This would enable us to take the four decades of research by Thorndike and his students, and gratefully utilize them to the full, simply noting that it was probably not the peripheral act, but the central process that was rewarded and punished in all these studies. In the same way we can take the massive materials of Hull and Skinner and gratefully note that they can without violence be interpreted as reinforcement of perception. Indeed, I suspect that Neal Miller's experiments will soon be leading to the same conclusion. For example, he and his associates (21) have directly excited affec-

tivity through electrical stimulation of the brain stem, apparently finding that this serves very well as a motive. From this point it is only one step to the point in which some types of brain stem stimulation may be used as "rewards" without any peripheral tension reduction.

We are working here with a very simple form of perceptual learning. There are others. We make no squatter claims. With Tolman (31) we are sure that "there is more than one kind of learning." But we suspect that that garden variety of learning that offers the classical basis for our vast system of social rewards and punishments is a matter of instilling in the individual a new way of perceiving the environment. Some things, originally neutral, he comes to see as desirable; others, originally neutral, he comes to see as undesirable. The affect may range from abject horror to the mystic's sense of unutterable joys, and the connections formed may range from the simplest nonsense-syllable kind of connection to the profoundest creative integration. Through all this, I believe, runs a central theme: the perceptual field comes to take on a structure in which the acceptable, the good, the satisfying, tends to take the dominant position. Some will think that this is too lowly, too primitive a view of human nature. I cannot see it that way. Socrates in the dialogue of the Gorgias said he was persuaded that all men alike seek joy, and that the difference between a noble man and a base man lay not in the operation of this simple principle, but in the wisdom which showed the noble man that virtue brought the greatest joy. It is not by rejecting man's ineradicable earthy quest of joy, but by understanding it and its consequences that the science of man may be helpful to man.

REFERENCES

1. AYLLON, T., & SOMMER, R. Autism, emphasis, and figure-ground perception. *J. Psychol.*, in press.
2. BIRCH, H. E., & BITTERMAN, M. E. Sensory integration and cognitive theory. *Psychol. Rev.*, 1951, **58**, 355–361.
3. DULANY, D. E., JR. Avoidance learning of perceptual defense and vigilance. Unpublished doctor's dissertation, Univer. of Michigan, 1954.
4. GIBSON, J. J., & GIBSON, ELEANOR J. Perceptual learning: differentiation or enrichment? *Psychol. Rev.*, 1955, **62**, 32–41.
5. HARLOW, H. F., HARLOW, MARGARET K., & MEYER, D. R. Learning motivated by a manipulation drive. *J. exp. Psychol.*, 1950, **40**, 228–234.
6. HOLT, E. B. *Animal drive and the learning process.* New York: Holt, 1931.
7. HULL, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
8. JACKSON, D. N. A further examination of the role of autism in a visual figure-ground relationship. *J. Psychol.*, 1954, **38**, 339–357.
9. KILPATRICK, F. P. *Human behavior from the transactional viewpoint.* Hanover: Inst. for Assoc. Research, 1952.
10. LEVINE, R., CHEIN, I., & MURPHY, G. The relation of the intensity of a need to the amount of perceptual distortion: a preliminary report. *J. Psychol.*, 1942, **13**, 282–293.
11. MELTZER, H. The present status of experimental studies on the relationship of feeling to memory. *Psychol. Rev.*, 1930, **37**, 124–139.
12. MELTZER, H. Individual differences in forgetting pleasant and unpleasant experiences. *J. educ. Psychol.*, 1930, **21**, 399–409.
13. MUENZINGER, K. F. Motivation in learning: I. Electric shock for correct response in the visual discrimination habit. *J. comp. Psychol.*, 1934, **17**, 267–277.
14. MURRAY, H. A. *Explorations in personality.* New York: Oxford, 1938.
15. POSTMAN, L. Association theory and perceptual learning. *Psychol. Rev.*, 1955, **62**, 438–446.
16. POSTMAN, L., & MURPHY, G. The factor of attitude in associative memory. *J. exp. Psychol.*, 1943, **33**, 228–238.

17. PUSTELL, E. Cue and drive aspects of anxiety in relation to perceptual vigilance and defense. Unpublished doctor's dissertation, Univer. of Michigan, 1954.
18. RAZRAN, G. H. S. Conditioned responses: an experimental study and a theoretical analysis. *Arch. Psychol.*, 1935, No. 191.
19. RAZRAN, G. H. S. Attitudinal control of human conditioning. *J. Psychol.*, 1936, **2**, 327–337.
20. RAZRAN, G. H. S. Conditioning and perception. *Psychol. Rev.*, 1955, **62**, 83–95.
21. ROBERTS, W. W., MILLER, N. E., & DELGADO, J. M. R. Motivation of learning by electrical stimulation in the diencephalon. *Amer. Psychologist*, 1954, **9**, 456–457. (Abstract)
22. ROCK, I., & FLECK, F. S. A re-examination of the effect of monetary reward and punishment in figure-ground perception. *J. exp. Psychol.*, 1950, **40**, 766–776.
23. RUBIN, E. Visuell Wahrgenommene Figuren. Copenhagen: Gyldendolska, 1921.
24. SANFORD, R. N. The effects of abstinence from food upon imaginal processes: a preliminary experiment. *J. Psychol.*, 1937, **3**, 145–159.
25. SCHAFER, R., & MURPHY, G. The role of autism in a visual figure-ground relationship. *J. exp. Psychol.*, 1943, **32**, 335–343.
26. SIMPSON, J. E. The influence of auditory stimulation on aniseikonic perception. *J. Psychol.*, in press.
27. SKINNER, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.
28. SMITH, D. E., & HOCHBERG, J. E. The effect of "punishment" (electric shock) on figure-ground perception. *J. Psychol.*, 1954, **38**, 83–87.
29. SNYDER, F., & SNYDER, C. The effect of monetary reward and punishment on auditory perception. *J. Psychol.*, in press.
30. TOLMAN, E. C. *Purposive behavior in animals and men.* New York: Century, 1932.
31. TOLMAN, E. C. There is more than one kind of learning. *Psychol. Rev.*, 1949, **56**, 144–155.
32. WOODWORTH, R. S. Reënforcement of perception. *Amer. J. Psychol.*, 1947, **60**, 119–124.

(Received January 7, 1955)

Psychological Review
Vol. 63, No. 1, 1956

AN AXIOMATIC AND EXPERIMENTAL STUDY OF SENSORY ORDER AND MEASURE [1, 2, 3]

EUGENE H. GALANTER

University of Pennsylvania

I

It is difficult to detect color blindness in a person, since the content of his awareness is not and cannot be known. But we find that such detection is possible and positive when we obtain from such a color-blind person information concerning the structural or relational modes of his behavior. Clearly the distinctions in content which form the intuitive basis for our judgment concerning the anomalousness of the vision of the color blind do not serve as evidence for the existence of the anomaly. The *content* of the mind (or, periphrastically, the *meaning* of behavior), being forever unknowable, cannot serve as the basis for psychological science, even though it may subserve intuitive judgment.

In casting about for a foundational principle for the science we should first note that there are certain characteristics of experience that we can communicate. Thus the statement (true or false) that "red looks more like orange than green does" suggests that some aspects of the mind are observable through behaviorally identifiable discriminations. This communicated mentation is the structure or the relationship between the elements of sensation, and not the elements themselves (12).

Many psychologists believe that specific aspects of the physical environment communicate themselves to organisms, in the sense that the organism comes in contact with some invariant property or quality, exuded or projected from the complex of energy which goes by the name of "stimulus." It is often presupposed that for each stimulus there is its unique "quality" or "afferent stimulus event" or "cognitive process." But if we agree that content is not relevant in a discussion of psychology, then we must also agree that "stimuli" cannot actuate a process which represents them as "the content" of our consciousness. If we attempt to beg off by pleading "behavior only," we are left in the same muddle with respect to "responses." To wit, we would be forced to investigate the structural relationships between responses or response classes, which is tantamount to the problem posed by "stimuli."

This conception of psychology stresses the study of the order of the organism's experience as determined by a study of behavior. Note that it is the structural relationships of the behavior that we

[1] This study is a part of an interdisciplinary project of the Graduate School of Arts and Sciences of the University of Pennsylvania entitled: "A Basic Investigation of Order and Measure." Besides the author, the investigators include Professors Nelson Goodman (philosophy), N. J. Fine (mathematics), and G. E. Schweigert (mathematics).

[2] Some of the experimental work reported in this paper was aided by the Social Science Research Council through an Undergraduate Research Grant for the Summer of 1954 to Mr. Stephen Fox. Mr. Fox served as the experimenter for all the data collected to date. Much of the technical aid with respect to the apparatus was obtained from Professor W. A. Shaw, who also gave generously of time and patience in discussions on sensory psychology. Additional support for the research has been received from the Committee for the Advancement of Research of the University of Pennsylvania.

[3] Portions of this paper were presented at the 1954 meetings of the American Psychological Association, New York, under the title "An Axiomatic Theory of Sensory Orders."

assign as the subject matter for psychological research. To investigate such structures we are forced to consider what is undoubtedly one of the most important aspects of any science, the notion of order and measurement. Research into the logical foundations of the general problem of ordering and measurement first came into prominence around the turn of the century; it reached a rather stable condition in its original formalization for the physical sciences by Campbell, with the publication of his book, *Physics, the Elements* (6), and was dealt with again by Bridgman in *The Logic of Modern Physics* (5).

The original investigations upon which this paper is based were begun by Goodman in 1930. As Boring's student he had become aware of the problems of order and measure in psychology, and especially in the formal aspects of these problems. His research proceeded from that time until its culmination in the publication of his book in 1951, *The Structure of Appearance* (9). Two chapters in this book deal with order and measure, and present a new basis for the construction of sensory orders. To comprehend the ramifications of the Goodman development, it is first worth while to examine more closely the historical texture of the problem.

One suggested route is classical psychophysics, where researchers have set for themselves a problem that can be formalized in the following way. On introspection we find a set of experiences displaying some order, both qualitative and quantitative, that somehow reflects an orderliness external to us. Second, as physicists we find a distribution of energies that can also be ordered and interrelated. We also notice, if we are careful (2), that the orders are not one-to-one. Third, we presume a *rational* or *intelligible* function that maps values of the physical order onto or into our experiences.

To define any function there are three things to consider. First is the domain of one variable; for the psychophysicist the values of this variable correspond to certain units like millilumberts, lamberts, sound pressure, etc. Next, we order the range of values for our dependent variable, i.e., experience, as reflected in discriminative responses. Finally we find the rule or function that associates every element in the domain with some element in the range. In general, the domain and range of a function, considered as abstract sets, do not have to have any ordering imposed upon them. But if we want a function to be well behaved, to move by small steps in the range when small steps are taken in the domain, either we allow the function to impose an order on the sets, or we order the sets in advance. The psychophysicist normally allows the physical energy order and the preconceived order of experience to define the rule or function. Sometimes the rule is derived a priori, in which case the order of experience emerges. In both cases it is clear that the absence of at least two of the three components of the function leads to the use of intuitive plausibility as the *method* for discovering the order.

Thus if we know the order of the physical energy values, as we presumably do, it is necessary that we know either the function which translates these values into experience, or the order of experience. Fechner presumed that he was aware of the order of physical magnitudes and also of the experience order. Other, more modern, psychophysicists also believe that they are cognizant of the order of experience, or at least that their subjects are, and try to find the rule that relates the experiential dimensions to physical parameters. Generally these psychophysicists believe that subjects are able to bisect

experience along certain "dimensions," or that the subject knows what "brighter than" or "bluer than" means.

Although most of the research in psychophysics has been concerned with the establishment of correlations between physical parameters and discriminability, some attempts at order representation have been made. The order of color experience as depicted by the three-dimensional color solid is one such attempt at map making. The empirical difficulties with this construction suggest that this map is only a rough approximation and possibly no approximation at all, as evidenced by Boring's recent attempt (4) to construct a four-dimensional solid. Helson (10) has proposed that visual experience can be conceived as varying along 16 possible dimensions. Certainly, if he is correct, the tri- or even tetradimensional color solid is woefully underdimensioned. Henning's (13) odor prism is another such ordering scheme which replaces experimentation with rationalism. The difficulty with all these constructions is that the order is applied to the experience without a formalization of the method employed for generating the order.

But how can experimentation be carried out on the construction of orders? We cannot ask subjects to judge "greater than" or "less than," since we do not know what they are judging as "greater" or "less." In addition, the use of such an asymmetrical relation, although apparently necessary to establish orders, requires that for each new dimension that we propose we introduce a new relation, e.g., "is prettier than," "is colder than," or "is more neutral than." Each one of these relations establishes a new dimension, but observe that the new dimension was *not* brought into being by the perceiver whose experiences we are trying to map, but by the experimenter, whose intuitions and insights should be applied to *describing* the data, not *constructing* them.

Perhaps we could use some such relation as "is just noticeably different from." However, two qualities are just noticeably different only if in some *presupposed* order there is no element between them that is noticeably different from either; e.g., we find this predicate being used to determine the difference limen for hue (11). Here we are assuming in advance that the order of experienced qualities is the same as some order of stimuli, in this case wave length or frequency. To set out in this way is certainly to beg the question. Clearly, the linear order generated by this technique is far from our presystematic intuition of a circular order of hues. We must point out that here we are not questioning the value of determining the minimum change in wave length necessary for a discrimination, but rather the logical difficulty in using such a relation for the construction of orders.

A second criticism of the use of just noticeable differences emerges from a consideration of its logical properties. Any consistent use of the term implies symmetry between the components of the relation. Thus if a is just noticeably different from b, we presume that b is also just noticeably different from a. Now this may or may not be the case experimentally, as we shall point out later. If symmetry happens not to be the case, we have no method of correcting the peculiarity, and we are forced into redefining the term or junking the method. This difficulty is of course a special case of the more general problem cited above, namely, that any use of the jnd presupposes the existence of an order underlying the experience, which is isomorphic to the physical scale being manipulated by the experimenter.

Finally we might suggest the use of some similarity predicate, as suggested

by Richardson (14). In using the triad method we avoid the error of prejudging the dimensionality of the ordering. However, when we use this relation we find that we are dealing with a very complex predicate, since it must read "x is more similar to y than it is to z." This means that in any experimental situation the number of discrete judgments required of the subject becomes astronomical, as Abelson (1) has pointed out. In the case of 16 stimuli (which is the number we are presently using) the number of triads approximates 3300. Attempts to shortcut this difficulty have led Abelson to require of his subjects that they judge "psychological distance" between stimuli on some subjective scale. This, of course, reintroduces all the subjectivity we have been trying to undercut. Such judgments are clearly more complex than triad judgments, and their formalization would be hounded by difficulties that even triad methods have not yet solved.

It is also worth noting that even for the triad method no complete formalization has been forthcoming. Torgerson (17) has provided the most detailed analysis of the technique to date. His method yields computational formulae for constructing a distance matrix which can be mapped onto a Euclidean space. The orientation of the obtained dimensions, however, is arbitrary, and requires some criterion such as "simple structure" to provide a unique solution. The work of Coombs (7) and the Stanford Value Theory Project (16) are promising, although the construction of purely linear orders is their major concern. Clearly much groundwork remains to be done if triad techniques are to provide solid solutions to our problem. Any calculus which we are to find suitable should be completely formalized and abstract, so that it can be of use in all areas of psychology, including attitude scaling, psychiatric diagnosis, and group-structure analysis.

From the previous discussion it becomes apparent that one major theoretical task is to provide a logical apparatus capable of analyzing the experience order. Our sojourn into psychophysics and scaling theory reveals that this area is not yet ready to provide an answer. We find ourselves in the position of having to determine the structural relationships of our experience when we have no advance knowledge of the form that a map of these relationships may take. If we could approximate such a map, further empirical research would indicate to us where we could stretch forward for further data, and where we would need to revamp our ideas concerning those aspects of the world which are important determiners of response probabilities. When we state that we have no advance knowledge of the form which such a map would take, this means that we have no conception of its dimensionality, nor of the units of measurement that would be involved in its utilization, etc.

The central problem is to construct for each sense realm a map that will assign to each element in the realm a unique position, and that will represent relative likeness of experience by relative nearness on the map. Then, by defining a set of coordinates, we can develop a systematic nomenclature for the elements, and define specific color, and measure notions. It would then be possible to perform a preliminary test of the construction by examining whether the map does any gross injustice, like placing red next to green. The success of later definitions, like size and measure, would constitute the ultimate test of the order. We would beg the issue to test the map only by its correspondence with the color solid, for example, since this itself is an attempt at map making, and we are interested in check-

ing *its* validity also. Finally, once the map is constructed it will serve as the criterion in those difficult cases where intuition gives no guidance (9).

The first thought in attempting such a map-making project would be to examine those areas of mathematics which deal with orders. When we do so we find that in the construction of partly ordered sets, the basic inclusion relation is antisymmetric and transitive; this finding, if our previous discussions concerning dimensionality are correct, will not be of much help with the beginnings of our project. In addition, the construction of a lattice representing experience does not provide the characteristics of uniqueness which we feel a map must have. Lattice theory will be required in the later stages of the development of sensory scales, but the first approximate mappings cannot be obtained by using the apparatus of partial orders.[4]

[4] The author, in collaboration with Schweigert, has attempted to reformulate the logical apparatus presented in this paper by translating the primitive notions and definitions into a lattice theoretic structure. By identifying the primitive relation "matching" with the intersection of subsets of some universe, and equating the operation "†" (cf. p. 14 f.) with a subgroup operation on the algebra, we were able to convert the universe to a commutative Boolean ring. Having defined the group operation on this ring, it then seemed possible to consider the order imposed on the universe by the group operation. But since "intersection" generalizes over all subsets, every subset intersects with the universe, rendering an order impossible. We might save the construction by imposing certain restrictions on the formation of subsets and their intersections. However, the total number of restrictions necessary to effect an ordering would be so unwieldly as to make the technique worthless. From the preceding considerations we have come to the conclusion that available mathematical methods are not *obviously* applicable to the problem which we face.

II

As we pointed out previously, a formal system for order and measure has been forthcoming (9). It will be our purpose to describe this method with some modifications and notational changes. Finally we will put the technique to work on some preliminary experimental findings that have an intrinsic interest for sensory psychologists. We will then suggest problems that we have run into and that have not yet been solved.

Instead of beginning with the assumption that any ordering relation must be nonsymmetrical and transitive, we utilize the fact that elements which appear phenomenally identical may be the result of two different energy configurations (3, 15). Clearly, if we have three different stimulus events, say a, b, and c, then we justifiably regard b as being between a and c if the latter two do not match while b matches both. The underlying principles behind this judgment suggest the utility of "matching" as a basic relation for order. As a matter of fact, this judgment forms the basic rule for order, which will later be formalized as an "extrasystematic ordering rule" for the system. This predicate, "matching," from now on symbolized M,[5] holds between two elements of experience that on

[5] The logical notation used in this paper is translated below.

(x) = the universal quantifier, "for all x."
$(\exists x)$ = the existential quantifier, "for some x."
M = the basic primitive relation, "matches."
ψ = the basic primitive predicate, "is a color."
$\dagger$ = symmetric difference.
$>$ = "cardinally greater."
$<$ = "part of" or "subset."
$\ll$ = "proper part" or "proper subset."
$- -$ = "not."
$\cdot$ = "and."
$\rightarrow$ = "if . . . then . . ."
$\rightarrow\!\!\!\rightarrow$ = "if and only if."
$+$ = logical sum.
$-$ = logical difference.
$=$ = "is equal to."

immediate inspection are not discriminable. It is a two-argument relation, which makes it (theoretically) less complicated than "similarity," and provides relief from the inordinate number of judgments required for a triadic predicate. It is obviously applicable to all categories of experience, so that no imposition is made by the theory with respect to the kinds of data that can be analyzed. M is symmetrical, reflexive, but not transitive; i.e., it is not an equivalence relation. The statement that

$$xMy .\rightarrow. x = y \qquad [1]$$

will be a false statement for this system (except for the trivial case of a one-element universe); while the statement

$$(z)(zMx \rightarrow\!\!\!\rightarrow zMy) .\rightarrow. (x = y) \qquad [2]$$

will be true for this system. An element then not only matches itself, but may also match other elements without being identical to them.

To avoid philosophical issues which have no place in the discussion of the logical theory, we have restated Goodman's nominalistic definitions and axioms in terms of a finite Boolean algebra. This restatement identifies an element in the interpretation of the theory, say a "color," with the set or class of phenomenal appearances of a particular physical energy configuration. The elements of the system could just as easily be conceived as probability distributions of certain responses to the configurations. In that case, elements of the system would match if the difference between moments of the distributions did not reach some arbitrarily (perhaps physiologically) determined value. Considered as sets of stimuli, matching obtains in those cases in which, on some selected fraction of trials, pairs selected from the sets in question are not discriminable.

We will need for the development of the theory some principles of matching. No rigorous distinction will be made between axioms or theorems; however, definitions will be indicated by a "D" following the formula. Matching applies between all sets of phenomenal colors, sounds, positions, etc., and only to such sets.

$$\psi(x) \rightarrow\!\!\!\rightarrow (\exists y)(xMy). \qquad [3]$$

It is symmetrical and reflexive

$$xMy \rightarrow\!\!\!\rightarrow yMx \qquad [4]$$

$$\psi(x) \rightarrow xMx \qquad [5]$$

but not transitive

$$--(xMy \,.\, yMz .\rightarrow. xMz). \qquad [6]$$

The relationship between matching and identity can now be summarized in a single theorem

$$\psi(x) .\rightarrow: (x = y) .\rightarrow\!\!\!\rightarrow. \quad (z)(zMx \rightarrow\!\!\!\rightarrow zMy). \qquad [7]$$

It will often be necessary to speak about that class of sets all of which match a particular set or element. Such a collection will be called the "manor" of the set in question. Thus we can define the "manor of x":

$$\text{the manor of } x = {}_{df}(\imath y)$$
$$[(z)(xMz \rightarrow\!\!\!\rightarrow \psi(z) \,.\, z \ll y)]. \qquad [D8]$$

Or, stated in the natural language, the manor of x is defined as that unique class y such that, for all z, x matches z if and only if z is an element of the system, and is also a proper subset of y. Rather than introduce a new symbol for this, we will adopt the following notational rule:

Rule: Any expression consisting of "the manor of" followed by a lower case letter may be entirely replaced by the corresponding capital letter. [9]

Thus a shorter equivalent of [7] is

$$\psi(x) .\rightarrow. (x = y) \rightarrow\!\!\!\rightarrow (X = Y). \qquad [7a]$$

Every set like x and nothing else has a manor

$$\psi(x) .\rightarrow\!\!\!\rightarrow. (\exists y)(y = X). \qquad [10]$$

Thus unless some set x exists, X is an empty description. Every set is a proper part of its own manor and the manor of any matching set

$$\psi(x) \rightarrow (x \ll X) \qquad [11]$$

$$xMy \rightarrow (x \ll Y). \qquad [12]$$

And in addition every proper subset of the manor of x matches x

$$\psi(y) \,.\, (y \ll X) \,.\!\rightarrow\!. \, xMy. \qquad [13]$$

It will often be necessary to refer to the sum of the differences of the manors of two sets. In order to do this we will adopt the following abbreviation:

$$X \dagger Y =_{df} (X - Y) + (Y - X), \quad [D14]$$

which is equivalent to

$$X \dagger Y =_{df} (X + Y) - (XY), \quad [D15]$$

and the condition for existence can be stated as

$$(\exists z)(z = X \dagger Y) \\ .\!\rightarrow\!. \psi(x) \,.\, \psi(y) \,.\, (x \neq y). \qquad [16]$$

Obviously "$\dagger$" is commutative:

$$X \dagger Y = Y \dagger X. \qquad [17]$$

We will now display, for a simple linear order, quite briefly and without proof (although proof is available), how the preceding calculus can be applied to a set of data collected by having the subjects match stimuli. The method for collecting such data is not yet systematically developed. We are now using a modification of the method of paired comparisons. Other methods are being tested, and although these problems offer tremendous experimental difficulties, the logical technique is immune to the experimental confusion.

First, we specify the single *extra-systematic* rule for order which we will employ, namely, that *the span or distance between any two matching elements (colors, attitudes, opinions, syndromes, people, etc.) is less than the span between any two nonmatching elements*. Note that this is simply a restatement of our original intuition concerning the utility of matching as an ordering predicate.

Let the first 16 letters of the alphabet be the collection of colors to be ordered, and let our data indicate that:

a matches *a,b,c,d*
b matches *a,b,c,d,e*
c matches *a,b,c,d,e,f*
d matches *a,b,c,d,e,f,g*
e matches *b,c,d,e,f,g,h*
f matches *c,d,e,f,g,h,i*
g matches *d,e,f,g,h,i,j*
h matches *e,f,g,h,i,j,k*
etc.

Also, each letter in our data matches only those letters not more than three places removed from it. The first notion we need to define is "betwixtness," which tells us when any element, say y, is between x and z, which we will note as $x/y/z$.

$$x/y/z =_{df} xMy \,.\, xMz \,.\, yMz \\ .\, (X \dagger Z > X \dagger Y) \\ .\, (X \dagger Z > Y \dagger Z), \quad [D18]$$

where "$>$" is to be read: "is cardinally greater than." We can show that in the example d is between c and e by substitution in the definition as follows:

$$c/d/e :\!\rightarrow\!\rightarrow\!: \frac{(cMd \,.\, cMe \,.\, dMe)}{\text{true}}$$

$$\cdot \frac{\dfrac{(C\dagger E > C\dagger D)}{\begin{matrix}(ahg) & (g)\\ 3 & > \ 1\end{matrix}} \quad \text{and} \quad \dfrac{(C\dagger E > D\dagger E)}{\begin{matrix}(ahg) & (ah)\\ 3 & > \ 2\end{matrix}}}{\text{true}}$$

[D18] with other theorems of the calculus yields some interesting derivations, e.g.,

$$--(x/x/y), \qquad [19]$$

$$--(y/x/x), \qquad [20]$$

$$--(x/y/x), \qquad [21]$$

$$x/y/z \,.\!\rightarrow\!. \, x \neq y \,.\, x \neq z \,.\, y \neq z, \qquad [22]$$

$$x/y/z .\rightarrow. z/y/x, \quad [23]$$

$$x/y/z .\rightarrow. --(x/z/y) . --(y/x/z). \quad [24]$$

[D18] along with [19] to [24] specify the conditions under which one color will be between any other two.

However, to establish an order on this set we must proceed to a definition of besideness. It is this relation which will allow the reconstruction of *a,b,c,d,e,f,g, h,i,j,k,l,m,n,o,p* from our data. We define "beside" as a two-place predicate obtaining between sets as follows:

$$xBy =_{df} xMy . x \neq y . (z)--(x/z/y), \quad [D25]$$

and a few theorems

$$--(xBx), \quad [26]$$

$$xBy .\rightarrow. yBx, \quad [27]$$

but B is nontransitive.

With the definition of "beside" it is now possible to reconstruct the original ordering. It is interesting to notice that the notion of a "linear" array can *almost* be formally specified by the two following characteristics:

1. having not more than two elements beside any one given element; and
2. having two elements such that only one other element is beside each, i.e., having two ends.

But notice that orders consisting of two discrete circular orders and one discrete linear order also fulfill the conditions of the formal specification for "linear" array. Obviously we want our idea of an array to contain some continuity or connectedness principle. Such a notion is also a necessity if we are ever to divide the sensory array into distinct sense realms, like "vision" and "audition." The concept we need must correspond to some notion similar to that of "connectedness" in graph theory. To do this within the framework of our apparatus we define a new kind of set which we term a "clan":

$$Cx =_{df} (y)[y \ll x .\rightarrow. (\exists r)(\exists s) (r < y . s < (x - y) . rMs)]. \quad [28]$$

This definition points out that the set x is a clan if, for every proper subset y, there exist two distinct subsets, r and s, where r is in y and s is in x minus y, and r and s match. Now it can be seen that a clan is a set which contains an M path between each two of its parts, and thus fulfills the requirement we pointed out above. Thus the notion of an "array" refers only to clans. It is to clans then that "linear" applies.

Such orders are best called "open linear orders" or "chains," to distinguish them from linear orders which have no ends, e.g., "circular linear orders," or from "cyclical linear orders." The example which we employed was purposely kept uncomplicated to demonstrate how the relation M can be used to develop an order on some set of stimuli.

The next notion with which any ordering theory must deal is the concept of "measurement." Once an order is established it is convenient to mark off certain units along its various "dimensions" which have properties usually associated with measurement. One of the most common concepts in psychology for the performance of this task is the "just noticeable difference." With some linear array obtained by the previous technique, it would seem an easy matter to determine which elements were just noticeably different from each other. Roughly speaking, if x is an element in a linear array, then the first element in either direction from x which does not match x may be said to be just noticeably different. The trouble is that, when we use this criterion, "just noticeably different" unexpectedly turns out to be nonsymmetrical! The following example illustrates this point. Suppose we have a set

of elements as follows:

$$\begin{aligned} A &= a,b \\ B &= a,b,c,d \\ C &= \quad b,c,d \\ D &= \quad b,c,d,e \\ E &= \quad\quad d,e,f \\ F &= \quad\quad e,f \end{aligned}$$

The resultant linear order appears to be $a,b,c,d\ e,f$. Now observe that the first letter to the right of b that does not match b is e. But the first letter to the left of e that does not match it is no longer b but rather c. No matter what is attempted an anomaly results. It seems impossible, with such an ordering, to establish an adequate concept of just noticeable difference. But observe that such an array does not satisfy our rule for ordering, i.e., that the span between any two matching elements is *less* than the span between any two nonmatching elements. So we are faced with the problem of constructing, from such data, maps that satisfy this rule. Since the adequacy of the definition of betwixtness determines the order of the arrangement, no rearrangement of these elements will work, for every other order must also violate the rule. Thus without changing the order already established by the definitions we must somehow construct maps that meet this requirement. We will call arrays such as the above "irregular," and define an irregular linear array as an array within which there are pairs of nonmatching elements with as many or fewer elements between them as there are between some two matching elements.

Fine (8) has proved an important theorem in which the notion of a "regular" array, and an array for which "just noticeably different" is symmetrical, are equivalent. This theorem clears up the problem we discussed earlier, since it provides a technique determined by the apparatus for theoretically correcting any asymmetry in jnd's. The fact that irregularity and asymmetry of jnd's coincide strongly suggests that the measurement problem is dependent upon a method for regularization for its solution.

But how are such irregularities to be removed? No technique of mapping will alter the fact that between the nonmatching elements e and c there are as many elements as between the matching elements b and d. Obviously the rule of order is violated. But in plotting the array on paper we can compensate for these irregularities by making appropriate adjustments in the spacing. For example, by incorporating in our map certain positions at which we plot no elements, the array can be regularized. By judiciously selecting the positions to be left vacant we can satisfy our rule and eliminate the anomalies of "just noticeable differences."

The following set of eleven elements can be used as an example:

$$\begin{aligned} A &= a,b,c,d \\ B &= a,b,c,d,e \\ C &= a,b,c,d,e,f \\ D &= a,b,c,d,e,f,g \\ E &= \quad b\ c,d,e,f,g \\ F &= \quad\quad c,d,e,f,g,h \\ G &= \quad\quad\quad d,e,f,g,h,i \\ H &= \quad\quad\quad\quad f,g,h,i,j \\ I &= \quad\quad\quad\quad\quad g,h,i,j,k \\ J &= \quad\quad\quad\quad\quad\quad h,i,j,k \\ K &= \quad\quad\quad\quad\quad\quad\quad i,j,k \end{aligned}$$

This set yields the unadjusted map a,b,c,d,e,f,g,h,i,j,k. We need to interpolate a vacant position, v_1, between g and h; but that alone will not give us a fully adjusted map, since between the matching elements d and g there are as many map positions (two) as between the nonmatching elements h and k. We cannot put the new vacant position between h and i, for then we have three map positions (v_1, h, v_2) between the matching elements g and i. Thus v_2 must be put between i and j or between j and k. In either case the rule is satisfied. Yet

there is an independent reason for choosing the second alternative. If we put v_2 between i and j, then there will be four map positions between the just noticeably different elements g and j, while there will be only three between each other two just noticeably different elements. If we place v_2 between j and k, then each two just noticeably different elements in the array will have exactly three map positions between them. Since a fixed span of just noticeable differences is a convenient feature for measurement, v_2 is best placed between j and k. We have adopted a convention for speaking about the number of map positions which one can traverse and still find matching elements. We call this the "M span" of the map.

The distance between the elements in an array is most appropriately measured by the number of map positions between the two on an adjusted map. The distance between just noticeably different elements in a linear array will then be constant throughout, even though the number of elements between them varies. Of course, all of these preceding comments about "map adjustment" and so on are rather loose and imprecise. The entire theory of map adjustment is yet to be worked out, especially in the multidimensional case.

However, work in this direction is under way, and Fine has recently proved that every linear array can be embedded in a regularized linear array. Indeed, this embedding can be performed by means of an explicit algorithm, and an estimate can be found for the number of additional elements required for the regularization. It is not yet known whether the resulting extension is minimal, or whether the minimal extension is unique. These appear to be difficult problems.

So far we have been dealing with purely linear orders or chains. However, it is obviously the case that not all problems which we face will have linear maps. In many cases it will be impossible to construct an arrangement of elements satisfying the definition of linearity. For example, suppose we have the following elements:

$$A = d,a,b$$
$$B = \quad a,b,c$$
$$C = d \quad b,c$$
$$D = d,a \quad c$$

It is quite clear that no linear arrangement of the four letters will satisfy our rule. Plainly we need further cartographical conventions to accommodate this and many other cases. Such conventions have been developed for the simpler cases, but further research and discussion in this area are needed before we can even propose a meaningful definition of dimensionality.[6]

III

We have previously argued that psychophysical methods impose the orders which they aim to describe. The strength of our theory would most clearly reveal itself if experimental results could be marshalled which question seriously established views concerning the sensory order. We have at this time barely scratched the surface of the problem which we face experimentally. But since

[6] In the case where elements of a set are related on the basis of a binary predicate such as our "matching" predicate, the elements together with the matching lines form a one dimensional connected topological space. By a well-known theorem in topology, any space of n dimensions can be embedded or represented in a Euclidean space of $2n + 1$ dimensions. In our case this means that any array of elements connected by matching is representable in a three-dimensional diagram. The problem of "dimensionality," discussed above, should not be confused with this fact of "representation." As Fine has pointed out, it would be desirable to find an appropriate intrinsic definition of the dimensionality of an array, comparable to the one which now exists for "linear" (one-dimensional) arrays. We would like to be able to prove a theorem which states that every array of dimension n can be mapped into a regular lattice in n space, with a uniform M span. Such a definition is as yet unavailable.

we do have some data which are quite interesting, it seems worth while to sketch our findings, and to suggest the nature of the results to be reported in more detail at a later date.

We had subjects with normal color vision report "match" or "different" when presented with a pair of stimulus lights. The peak wave-length transmission of the lights could be varied over four values: 510, 507, 502, and 499 millimicrons. In addition, and simultaneously, the intensity of the lights could be set at one of three different levels, allowing a total of twelve different combinations of wave length and intensity. Each light could be independently adjusted so that each of the twelve colors could be compared with all of the others. The subjects were so trained that approximately 40 per cent of the judgments elicited a "match" response. This training was undertaken so that a sufficient density of "matching" responses would be obtained. (The effects of changes in this "criterion level" are yet to be investigated.) The subjects did not, of course, know the physical parameters which were being varied. They were told to report "same" or "different" with respect to the lights.

The physical values can be ordered in a two-dimensional array, as follows:

intensity		510	507	502	499
	hi	*a*	*d*	*g*	*j*
	me	*b*	*e*	*h*	*k*
	lo	*c*	*f*	*i*	*l*

From classical considerations in psychophysics we would expect that the experienced order would somehow parallel the physical order. We had the subjects make the comparisons, and, as a first approximation to the establishment of experimentally derived statistical criteria, we accepted as a "match" those pairs for which the subjects reported matching more than 50 per cent of the time.

The data which we obtained for all the subjects can be summarized in the table of manors listed below.[7]

$A = a,b,\ d,e$
$B = a,b,c,d,e,f$
$C = \ \ b,c,d,e,f$
$D = a,b,c,d,e$
$E = a,b,c,d,e,f,g$
$F = \ \ b,c\ \ e,f\ \ h,i$
$G = \ \ \ \ \ e\ \ g,h,i$
$H = \ \ \ \ \ \ f,g,h,i,j$
$I = \ \ \ \ \ \ f,g,h,i,j$
$J = \ \ \ \ \ \ \ \ h,i,j,k$
$K = \ \ \ \ \ \ \ \ \ \ j,k,l$
$L = \ \ \ \ \ \ \ \ \ \ \ \ k,l$

Note that h and i collapse to a single element (by [7a]) which we will name "*hi.*" Applying the techniques of the calculus to the data yields an irregular array as shown in Fig. 1, and, as regularized, in Fig. 2. Except for the g element, the order is essentially linear. With g deleted, the regularized array is as shown in Fig. 3. We suspect that some lack of statistical stability is responsible for the departure from linearity, and are presently collecting data for sixteen stimuli to examine this question.

We wish to stress that, in the absence of evidence with respect to uniqueness, it is impossible for us to argue for the ultimate validity of the representation, even assuming the statistical question to be unimportant. Using some criterion such as "minimum additional vacant spaces," we may argue that as a rough-and-ready method the present mapping

[7] Pooling data for three subjects must be seriously questioned in studies of this kind. These data are included merely to indicate the feasibility of the technique. Forthcoming empirical results will report individual data.

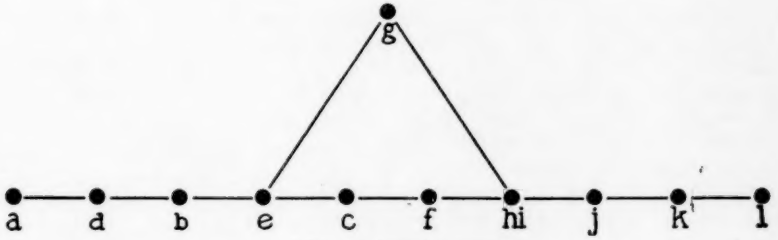

FIG. 1. Irregular array.

is in some sense the "correct" one. We have tried other representations, and have been unable to discover any other plan for mapping the elements which yields a constant *M* span, and which correctly reads off the manors.

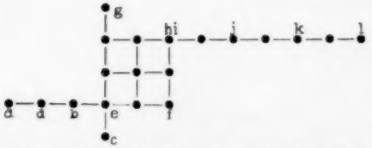

FIG. 2. Regularized array. The points without letters represent vacant positions. The *M* span equals 3; i.e., all elements within three steps of a given element match. The jnd therefore equals 4.

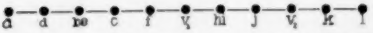

FIG. 3. Regularized array with *g* deleted. Note that with *g* removed, *b* and *e* collapse to a single element. The points labeled v_1 and v_2 represent vacant positions. *M* span equals 3; jnd equals 4.

The linear path through the two-dimensional physical continuum indicates that the subjects are not simply switching intensity after running through the four wave lengths or vice versa. No alteration of the physical order will support the conclusion that the subjects discriminate *two* qualities of the stimulus independently. Except for the maverick stimulus *g*, the order of the experience is linear to the subjects. However, care should be taken in the interpretation, since the quantity of data is still small.

Also being planned are experiments to test the measure notions of the theory. By obtaining an irregular linear array and its "unique" regularization, we can then introduce new stimuli which should be present for the vacant spaces. All other intermediate values should be identical to the elements already mapped. Thus we will be able to validate a principal notion with respect to measure, namely, that stimuli are present for all and only those map positions which can be regularly ordered. Theoretical and experimental research is also under way with respect to the representation of orders for color-anomalous individuals.

To demonstrate the breadth of the general apparatus, final mention should be made of some collateral research in the analysis of group structure. Since the apparatus is neutral with respect to the interpretation of the primitives, any binary relation satisfying the axioms of the system will yield orderings for the elements of the interpretation. We need only use a predicate which is symmetrical, reflexive, and nontransitive to relate individuals in a group, and this yields an ordering for the group. This ordering is descriptive of the group for the interpretation placed on the relation. We are at present examining group sociograms in which the basic datum represents the relations between individuals represented by the concept "very friendly with." Interpreting this relation as symmetrical and reflexive, and utilizing certain topological concepts developed in collaboration with Schweigert, we have been able

to characterize the compactness of the group, its clique structure, dominance, leadership, social distance, and other group notions. We have constructed a pseudometric for social distance by using the ratio of the cardinality of the symmetric difference between manors, normalized by the sum of the elements in both manors. For example, if *a* is very friendly with *a,b,c,d,e,f,* and *h* is very friendly with *d,e,f,g,h,i,j,k,l,* then the number of elements in the symmetric difference is 9, and the sum of the elements in both manors is 12, yielding a social distance of 9/12 or .75.

Finally, we might offer a word of comparison between the proposed theory and other more standard techniques. One disadvantage may be admitted: our method gives no *direct* ordering for a set of nonmatching elements; however, balancing this is the fact that other methods are often defeated when matching elements belong to the set to be ordered. Another feature which some may regard as a disadvantage is that at present we get a relative rather than an absolute order, although this is still a factual question and one which remains to be observed. On the other hand, the advantages offered by the calculus are (*a*) its generality for all sets of elements of all kinds; (*b*) the simplicity of the required response; (*c*) the neutrality of the method, which presupposes no prior order or analysis; (*d*) the fact that it provides a method for generating orders rather than merely checking consistency with some a priori order; and finally (*e*) that it is logically more explicit and complete than most methods at the present time.

SUMMARY

1. Various classical methods used for the construction of sensory orders are criticized, and criteria for a satisfactory technique are suggested.

2. A new method for the analysis of sensory order and measure is described. The formal characteristics of the technique are presented, together with illustrative examples.

3. Some data from color vision experiments are presented, and analyzed by the technique.

REFERENCES

1. ABELSON, R. P. A technique and a model for multidimensional scaling. *Pub. Opin. Quart.* In press.
2. BORING, E. G. The relation of the attributes of sensation to the dimensions of the stimulus. *Phil. Sci.*, 1935, **2**, 236–245.
3. BORING, E. G. The psychophysics of color tolerance. *Amer. J. Psychol.*, 1939, **52**, 383–448.
4. BORING, E. G. A color solid in four dimensions. *Année Psychol.*, 1951, **50**, 293–304.
5. BRIDGMAN, P. W. *The logic of modern physics.* New York: Macmillan, 1928.
6. CAMPBELL, N. R. *Physics: the elements.* Cambridge, England: Cambridge Univer. Press, 1920.
7. COOMBS, C. H. A theory of psychological scaling. *Univer. Mich. Engng. Res. Inst. Bull.*, 1951, No. 34, 94 pp.
8. FINE, N. J. Proof of a conjecture of Goodman. *J. symbol. Logic*, 1954, **19**, 41–44.
9. GOODMAN, N. *The structure of appearance.* Cambridge: Harvard Univer. Press, 1951 (esp. Chaps. IX and X).
10. HELSON, H. *Theoretical foundations of psychology.* New York: D. Van Nostrand, 1951.
11. JONES, L. A. The fundamental scale of pure hue and retinal sensitivity to hue differences. *J. Opt. Soc. Amer.*, 1917, **1**, 63–77.
12. LEWIS, C. I. *Mind and the world order.* New York: Scribner's, 1929.
13. MACDONALD, M. K. An experimental study of Henning's system of olfactory qualities. *Amer. J. Psychol.*, 1922, **33**, 535–553.
14. RICHARDSON, M. W. Multidimensional psychophysics. *Psychol. Bull.*, 1938, **35**, 659–660. (Abstract)
15. STEVENS, S. S. The attributes of tones. *Proc. Nat. Acad. Sci.*, 1934, **20**, 457–459.
16. SUPPES, P., & WINET, M. *Axiomatization and representations of difference structures.* Stanford Value Theory Project, Stanford Univer., 1954, Rep. No. 2.
17. TORGERSON, W. S. Multidimensional scaling: I. Theory and method. *Psychometrika*, 1952, **17**, 401–419.

(Received February 28, 1955)

Psychological Review
Vol. 63, No. 1, 1956

"FUNCTIONALISM" IN PERCEPTION [1]

W. C. H. PRENTICE

Swarthmore College

In the recent history of perceptual theory, what Bruner and Postman (4) have called "Functionalism" is clearly the major development. I shall make no attempt to define "Functionalism" beyond their own comments, though Brunswik (6), in the same Symposium, attempted to explicate the term by referring to several historically distinguishable versions of this approach to perception. Bruner and Postman had this to say:

> [This] approach attempts to treat perception in a broader behavioral context. For its primary concern is with the manner in which perceptual functioning is imbedded in and interacts with other forms of psychological functioning. Perception is viewed as instrumental activity. Those who take this approach to perception hope to add to and go beyond the variables and relationships discovered in the formal study of perception, not in any sense to disregard them. For a full understanding of the perceptual process it is necessary to vary not only the physical stimulus and the sensory state of the organism but also those central conditions—motives, predispositions, past learnings—which have largely remained outside the formal limits of the perceptual system. The data of such research—the dependent variables—are often the same as those of the Formalists: sensory discriminations and characteristics of perceptual organization (4, p. 15).

It is interesting that the typically functional American psychology almost ignored perception for so many years. This was perhaps because one can study the purposive and practical behavior of man without halting long over the perceptual processes that make such behavior possible. More complex processes of seeing and hearing can be thought of as idiosyncratic and left to students of learning or personality, while *general* laws of sensory experience are sought only for the kind of limiting responses studied in psychophysics.

In recent years, however, the view of perception as idiosyncratic, not only different in different people but correlated with emotional and motivational variation, began to show promise for a kind of "functional" interpretation. In this form, perception is conceived as a personal and goal-directed reaction, responsive to needs and attitudes, subject to training by success and failure, and forming a part of the uniqueness of each individual. In the absence of any detailed historical analysis, it seems fair to date the present enthusiasm for perceptual research from the paper by Bruner and Goodman (3) reporting a difference in the perceived size of coins for rich and for poor boys. That work and other papers from the same laboratory seem to have captured the imagination of a large number of workers. The general contention, you will remember, was that *need* is an important parameter in perceptual experiments and that what the subject sees or hears may be in large part determined by what he wants. It should be noted in passing that this has been an extraordinarily difficult thesis to establish because of the multiplicity of other influences that must be controlled, the Bruner-Goodman effect itself being extremely hard to reproduce. Nevertheless a rash of studies appeared connecting various kinds of need, attitude, social pressure and the like to perceiving.

A quite independent series of re-

[1] Based on a paper read at the Symposium on "Conceptual Trends in Psychology," American Psychological Association, New York, September, 1954.

searches, as influential as the Harvard ones, implies that we perceive what it is useful to perceive. These have been much less concerned with personal and motivational matters, and it is even possible to anticipate an ultimate clash between the two kinds of functional emphasis. If we see what we have learned to see, then being hungry might not affect our perception a bit unless certain very special kinds of learning opportunity had coincided with our being hungry in the past. The emphasis on experience also seems to suggest a more limited use of perceptual materials for personality analysis, for though both perception and other idiosyncrasies might depend on the past history of the individual, there is no obvious reason why they should be closely correlated. And the studies in question tend to stress the *generality* of the kind of experience that determines perception rather than its uniqueness.

It would be a mistake to assume that all students of functional and personal aspects of perception derive from one of these two predecessors. European psychologists and a few Americans had produced researches exemplifying what we would now call the functional approach. It is not possible to list all the experimental backgrounds of the current movement in an article of this kind, but the many little currents have swelled into a large river with contributions from many laboratories in many lands.

These groups taken all together have produced hundreds of research reports in the years since 1946. Together their findings constitute what might fairly be called "official" American psychology of perception today. It certainly has become the perceptual theory stressed by our introductory textbooks. And clinical psychologists have seized on perception as a possible index of personality, considerably extending the conceptions of Rorschach and of other pioneers in this field.

But these various kinds of "functionalism" are not without their opponents. There are those who hold that all this work hardly deals with perception at all. By saying that perception is a goal-directed response, they suggest, we merely inflate the meaning of the word perception to include far more than it means to traditional psychology (or to common sense). We run the risk of confusing the consequences of perceiving with the percept itself. If we interest ourselves in what people *see* instead of the remarks they make or the buttons they push, we may find that it is indeed possible to write general laws relating the physical stimulus to the percept.

Research on the psychophysics of space perception shows promise of doing, with different premises and different tools, what the gestalt workers set out to do thirty years ago—namely to rescue perception for general psychology. Considerable progress has been made toward specifying the stimulus conditions for color constancy—an event that seemed a few years ago to demand explanation in terms of unconscious inference. Still other studies have opened up the possibility that the aftereffects of prolonged stimulation may be the key with which to unlock certain mysteries of perception. Work on the Müller-Lyer illusion and its destruction by "training," for example, shows how stimulation-history, rather than learning in any usual sense, may account for both the illusion and its destruction. Even our understanding of social perception and of the perception of causal sequence is being attacked by analyses of stimulation.

The majority trend of "Functional" research has nevertheless brought us a long way toward accepting the view that everyone's perceptual world is different

and that what we see is pretty much of our own making. This conclusion is especially odd since we seem to know that autism is the exception and not the rule, that perceived color and size and shape are less often matters of personal idiosyncrasy than matters of lawful interaction betwen physical and biological systems. Certainly agreement among observers is more widespread than disagreement, and one must go to considerable trouble in the laboratory to establish situations that clearly produce disagreement in response to motives, personality differences, or past experience (e.g., **7**, **28**). We have, of course, learned that certain facts of perception can be understood only in terms of the life history of the particular perceiver, but we do not know how to make satisfactory generalizations about the facts in question. One cannot help being struck by the paucity of theories that might organize and systematize the relations between perception on the one hand and either need or experience on the other.

Perhaps our present confusion can be diminished by finding a common ground on which traditional or "formal" students of perception and contemporary "functional" workers can stand together, a ground sought by Bruner and Postman themselves (**4**), but still somehow eluding us. Perceptual theory is endangered by a failure to agree on the basic facts about experiments in perception. We should face the fact that "perceiving" is not something we observe directly and the further fact that rules for inferring about it are not generally agreed upon. In everyday life, when we want to know whether our subject can hear a given tone or see a color or a shape, we are likely to start by asking him: "What do you see (or hear)?" If he speaks the same language we do, and if he seems to understand the question, that may suffice. But sometimes it does not. Whether we accept his response as a useful one for our purposes depends on the nature of his answer. If it is ambiguous or startlingly different from that of other observers, or if we have reason to suspect the reliability of this kind of observation, we may persist by asking other questions or requiring the observer to make certain discriminations. If these additional facts still leave us in doubt about the observer's reliability or about our interpretation of his reactions, we may modify the conditions of his observation until we think we understand his total behavior and can infer what he is seeing. I want to stress the word *infer*. By "perceiving" we normally *mean* something by nature private, confined to the perceiver; all questions or tests lead merely to inferences on our part about what we take to be another's perceptions. We infer that our subject has perceived a particular stimulus pattern in a particular way only if his behavior makes any other explanation too improbable to accept. In the last analysis, each of us must (and does) decide for himself whether to be convinced. Discriminatory behavior may have a variety of explanations. Many people thought the famous horse, Clever Hans, could discriminate words and numbers because that seemed the only plausible explanation of his discriminatory behavior until Pfungst found the real one. Or take the case in which a subject *fails* to discriminate. If we are unable to teach an animal to push a bar to the right for a blue light and to the left for a red one, it may be because the lights look the same to him. But there are other possibilities, and it would be a serious error to say that what we *mean* by his seeing them as different is that he is able and willing to learn how an experimenter wishes him to respond.

In its usual sense, the term "perception" applies to that hypothetical event,

directly reportable (if at all) only by the subject himself, an event that may or may not modify any behavior observable to others. Unless we wish to appropriate an unfamiliar meaning for the term "perception," we should interpret changes in behavior as signaling changes in perception only when it has been established, beyond a reasonable doubt, that the changes have not resulted from some other response-selecting variable. Now, we can all agree that the techniques for arriving at that stage "beyond reasonable doubt" should in an ideal sense be spelled out exhaustively. We can even agree that if they were, they would in some sense constitute an operational definition of what we *mean* by perception. No one would claim, however, that in normal life or in the psychological laboratory we ever do so in fact.

This matter of usage can have important consequences. One widely quoted experiment (11) requires the subject to estimate the distance of a circular patch under special conditions. In one case the subject is told that the patch is a Ping-pong ball, for instance, and in another that it is a basketball. Subjects give differential responses, and the experimenters conclude that the perception of distance is affected by our past experience with meaningful objects. But one may ask: Does the location of the circular spot really *look* different when the instruction is changed? We cannot expect the investigators to answer that for us, of course, until we tell them what we mean by "really look different" in some kind of operational terms; and that is difficult to do. In practice one would try a number of things and finally come to a largely pragmatic or common-sense conclusion as to whether the complicated judgment about the "distance" of an ambiguously presented stimulus could *only* be explained by assuming a perceptual change.

Another example. The concept of "perceptual defense" somehow suggests that organisms can protect themselves by not perceiving certain threatening objects. How they discriminate the ones that would be dangerous to perceive is the question. Is it possible that the rejection is not a perceptual one but occurs at the level of report or other discriminatory *response?* Some psychologists tend to pass over such questions, being willing to accept a subject's different behavior in the face of threat words and control words as defining a perceptual difference—at least to a first approximation. Their critics implicitly use the behavioral evidence not as a definition of what was perceived but rather as grounds for an inference about what was perceived (cf. 16, 22, 25). I would urge that the latter is the more cautious course.

But Functionalists will grow impatient with this argument. "You seem to imply," I can hear them saying, "that some behavioral criteria are inherently more *perceptual* than others, that we should arbitrarily choose just some of the infinite variety of human modes of experience and response and call them perceptual. But the lines that distinguish perception from report, or perception from judgment, or perception from imagination are simply not clear. The artificiality of such distinctions is one of the cornerstones of the Functionalist position." In support of these statements, they can cite impressive evidence. We know that people may confuse imagination with veridical experience even under laboratory conditions. We have evidence for the influence of subliminal stimulation on guessing behavior; it has been shown that the observer may discriminate among tachistoscopically presented words on one level (e.g., GSR) but not on another (e.g., verbal report); both ancient and modern forms of the Stratton-type adaptation experi-

ment suggest that what the subject "sees" depends to some extent on what you ask him. And so on. No one need dispute these findings, and yet I wish that we could distinguish between perception and judgment, if only in the way that we distinguish between (e.g.) learning and maturation while acknowledging both the logical difficulties in making precise differentiations and the empirical difficulties in knowing which we are dealing with at any particular moment.

I am not yet prepared to admit that discriminatory behavior and perception are the same thing. Some years ago there was a considerable flurry of interest in intersensory phenomena, and psychologists found that subjects could make consistent judgments about such things as the lightness (on a black-white scale) of a tone or the weight (in grams or pounds) of a minute or an hour. It does not seem useful to me to call such responses perceptual. During the war I learned to estimate the distances of ships at sea fairly accurately after having been extremely inaccurate at first; when my reports of relative and absolute distances had changed remarkably, the ships still looked the same. I could recognize that there was a difference between my perception and my judgment of distance, and I keep feeling that a careful student of my behavior could have devised tools that would have let him make the same distinction.

We are really confronted, it seems to me, with a series of empirical questions about the *limits* of the traditional (or Formalist) view of perception and the Functionalist views. To what extent can perception be understood in terms of: (*a*) objectively measured aspects of the physical energies involved, (*b*) their relations to each other, (*c*) the duration and frequency of their occurrence, and (*d*) the structure, adaptation, and orientation of sense organs? To what extent can our failures of prediction from knowledge of such facts be traced to uncontrolled variables of instruction, understanding of the task, and the like? To what extent can variability of behavior that still remains be attributed to the training and past experience of the learner, and how does that learning become effective? Finally, to what extent must variability yet unaccounted for be attributed to emotional or motivational influences? We need a research program that is directed toward the answer to these questions in a systematic way.

If we conceive of these as the proper questions to ask, where do we now stand? *To what extent can generalizations be made about perception by trying to relate stimulus and response?* That depends on the kind of generalization we wish to make. In all psychophysical work, the variance due to individuals is rather high. In this sense, it is always difficult to predict the absolute response of an individual from group data. On the other hand, in many fields the group data themselves are remarkably stable, as are the data taken from a single individual. This is particularly true if we are concerned with formal relationships instead of with specific constants. Thus almost all human visual functions under darkroom conditions are stable and predictable. You and I can count on reproducing the textbook curves with a minimum of error. Further, initial intersubject variability can always be reduced to a remarkable degree by purifying set and instructions and by increasing the experience of the subject in making the kind of judgment required, thus suggesting that we were dealing primarily with response differences. Even so we cannot doubt that important individual differences remain. Every laboratory study of seeing confirms the long-familiar fact that human observers have different

visual constants. We simply do not know in many cases how much of these differences must be assigned to structural difference and how much to personal and experiential difference.

The best present answer to our first question seems to be that we can successfully make extensive generalizations about human visual and auditory perception as a function of physical and physiological variables, and that we can even state in most cases the limits within which individual differences will fall; in a few cases, moreover, we can state with confidence that the individual differences themselves reflect anatomical or physiological variation; in many more cases, of course, we know little or nothing about the source of the variability. Our confidence and precision get less the more complex the situation—that is, the larger the number of stimulus dimensions that can vary independently.

It is occasionally suggested that information theory can prove to be our salvation. Various writers (e.g., 20) have suggested that we define the stimulus in terms of response alternatives and that, once this is done, we may proceed to discover the limits within which variables of personality, set, motivation, etc., operate. This is an inviting notion and one that needs full exploration. Attneave (2) has made the boldest proposals in this respect. His work implies that perception may be conceived as a kind of guessing game, not unlike the reading of print. In fact, he uses the serial guessing technique that Shannon devised for language to show how visual figures may be generated by a series of binary choices. It is too soon to say whether the phenomena that appear in this highly artificial situation are typical of ordinary perception, but a comment or two may be in order.

First, the definition of response alternatives must itself be made on an empirical basis. Not until we have found the laws of perception that we are now seeking can we deduce the number of alternatives in a new stimulus array. Then what shall be the standard experimental condition under which the informational content of a stimulus pattern is defined? Are we not likely to re-live the whole Formalist-Functionalist controversy if we attempt to set up one "pure" standard with respect to which other situations will be judged? Perhaps some criterion of *maximum* information can be devised, but it is too soon to say.

Second, the information-theoretical approach makes it very tempting to define perceiving in ultraoperational terms, to write or speak as if the observer sees something different when we change his response possibilities. It has already become common for writers to point out how the number of responses permitted changes the distribution, and hence the relative accuracy, of certain kinds of discrimination. These facts have enormous importance for military and other applied problems where the accurate transmission of information is the criterion. It is clear that we can modify a transmission system that includes a human observer as one link. It is not clear that the responses permitted that observer to define uniquely the reservoir of experiences from which he draws his responses. Everyone knows that I may be asked to sort all the people I see into men and women and still recognize many differences among the members of each category. But when we are concerned with information transmission, we sometimes forget to try to study the maximum channel capacity under optimum conditions, or to wonder what happens to that capacity when we change the available response set.

These comments are warnings only. It may yet develop that information theory is indeed a powerful tool for the study of perception, but already there

are clouds on the horizon. Miller (23) reports cases of his own and of Pollack's in which certain kinds of grouping produce discrimination patterns that defy analysis in terms of simple binary choices and suggest the operation of "intervening variables" untouched by the communication-system model.

To what extent can these more complex perceptual situations be made to yield smaller interindividual variation? If, for example, we study size constancy in natural or representative settings (as Gibson [9] and Brunswik [5] and Gilinsky [10] have all done independently), to what degree can we reduce individual differences by changing nonperceptual aspects of the subject's behavior—e.g., by changing his mode of reporting, etc.? Once we have reduced them to a minimum, can we express the remaining differences by some kind of numerical constant, as Gilinsky found she could? Is there any reason to believe that these constants represent structural differences among our observers? Motivational differences? Differences of personality? Can we now increase the subject variance in systematic ways by modifying the experience or need structure of individual observers?

These questions particularly need to be asked because many studies (e.g., 1, 3, 7, 11, 17, 20) suggest strongly that qualitative or large quantitative differences among individuals can be demonstrated only in cases where the physical stimulation is inadequate to determine the response in any case, i.e., where unreliability of report is the rule rather than the exception. The experiments of Asch (1) are particularly revealing in this respect. His subjects were asked to make size judgments under conditions that would ordinarily have been completely stimulus-determined; they were subjected to extreme social pressures with the result that changes in verbal report, emotional response, and other aspects of behavior occurred in striking fashion. What did not change is what they *perceived;* inquiry makes it clear that only one or two subjects out of hundreds made any attempt to justify their changed reports by saying that the falsely reported line really looked equal to the standard. Other evidence makes it seem certain that these extremely rare protocols would become common in a test situation involving comparisons that are inherently difficult and productive of variable responses. We know how to produce such conditions. For vision we do it by presenting short, dim, out-of-focus exposures; we do it by asking for complex responses that are only in part determined by stimuli in any case—for instance, the reading of words; we do it by presenting stimulus situations in which the geometric or photometric conditions are known to lead to ambiguity —for instance, the traditional illusions or a heterochromatic brightness match. But suppose we reverse our aim. How *reliable* can we get our reports, using some one personal or motivational condition, and still be able to modify them by changing to another condition?

Unfortunately, we have almost no answers to these questions. Research has concentrated not on exploring the limits of effects but on maximizing favorable conditions. Some improvement in the situation may be hoped for as a result of explicit attention to stimulus variables in work like that of Lazarus, Yousem, and Arensberg (20). But we do not yet know whether our most dramatic demonstrations really have anything at all to do with everyday seeing.

When interindividual variation does persist beyond the stage of reasonable uniformity of treatment, how much of it can be attributed to learning and how much to temporary motivational condi-

tions operating directly on the perceiving mechanism? Most of the experiments on *need* are confounded in this respect, and it is hard to know whether the presence of the need merely permits the use of a different set of memories and thus a different response to the same percept, or whether the need really modifies the perceiving apparatus in some way. Postman (24) has commented at length on this problem, and Wallach, in particular, has demonstrated the very direct role of memorial processes in visual organization when some ambiguity is present (30, 31, 32). We do not have much information about the role of need when influences like Wallach's are ruled out. This continues to be one of the great gaps in perceptual research of the "functional" kind.

In those cases where we may have been able to establish an undoubted effect of past experience, how does the experience work? Ever since Helmholtz (and probably before), students of perception have written as if the perceptual mechanism makes *inferences*. Several writers with otherwise very different theoretical orientations have recently reaffirmed their confidence, for instance, in the concept of *cue* in perception. A cue is a sign or a signal; it sets in motion one train of activities rather than another. It results in the perceiver's consciously or unconsciously discriminating between the situation in which he should behave or see one way and the situation in which he should behave or see another way. Maybe some sorts of perceptual behavior do result from a kind of discrimination learning, but few if any instances of such occurrence have been unequivocally demonstrated. Can we, for instance, condition people to perceive in one way rather than in another? Several experiments of Gardner Murphy and his associates suggest that we can (21, 26, 29), but they are not immune to other possible interpretations. Various students (see especially 8) seem to have shown that synaesthetic or hallucinatory experiences may occasionally be induced by associational procedures. Studies with lower animals, like those on perceptual aspects of pecking behavior in chicks (14, 15), are equivocal but lend some support to the possibility that percepts may sometimes be formed in these ways. But these are isolated examples. Most of the areas in which we traditionally appeal to inference are ones in which learning has not yet been demonstrated to occur, and many examples strain the credulity of those who have studied learning in the laboratory. Possible alternative explanations are suggested by *adaptational* changes of the sort that have been demonstrated over short periods by such diverse methods as those of Helson (13) with color matching or weight lifting, those of Kohler (19) wearing filters or prisms over the eyes, or those of Köhler and Wallach (18) by staring at certain kinds of contour. We also know from the studies of Riesen (27) on chimpanzees reared in darkness that some conditions of early stimulation may severely limit or change the development of sensory equipment. Some of Hebb's (12) speculations suggest further ways in which experience or stimulation history may modify the percepts of organisms *without* associational or reinforcement learning in any usual sense. We know far too little about any of these processes and almost nothing about their limits. It is conceivable that their joint action, when understood, could remove much of the mystery (and all of the ghosts of unconscious inference) from our knowledge of perception.

To summarize: American psychology is and will undoubtedly continue to be a functional psychology. We tend to be interested in the behavior of the whole

organism and in the way experience or behavior fits into the economy or purpose of the living human being. For that reason, perception has not until recently been a markedly American subject of investigation. American textbooks have been notoriously inadequate in their accounts, and we have tended to think of perception as an uninteresting and unimportant side issue—if not, indeed, a blind alley. But for a variety of reasons which only a better historian than I can evaluate, American psychologists after World War II began to see some functional implications in the study of perception. They began to see how perception is in many cases imbedded in and formed by the activities and purposes and needs of the perceiver. They saw the possibility of throwing light on motivational and personal variables through the study of perception; they even saw perception as a kind of paradigm of all complex behavior. With the rush of investigations into perception as a function of posture, sex, hunger, past experience, and personality differences, it has been easy to forget that we do not know *how* these things happen. We do not seem to have acquired much theoretical insight by merely multiplying instances, and very little effort has been spent in studying the objective conditions underlying the effects themselves. It has been easy to forget or overlook the enormous number of cases that we know so well in which neither special interest nor training seems able to modify what we see, but rather in which the experience is quite precisely predictable from knowing physical facts about the stimulus and about the perceiver.

It is time to call for a return to the logic if not to the techniques of the psychophysicist. We run the risk of making of the field of perception what has been made of "extrasensory" perception: an area in which more and more demonstrations are carried out but in which theory and understanding never progress because certain kinds of systematic questions are not asked, because students are satisfied with analogies instead of with laws binding the phenomena to other known facts. So long as we are able to defend the use of so vague and analogical a word as *cue* in the field of perception, so long as we use terms that imply an *unbewusster Schluss* as an explanatory concept, so long as we do not distinguish between various kinds of learning in their effects on perception, so long as we do not attempt to explore the limits of the phenomena of need in perception or to inquire why they are so hard to demonstrate—just so long will we continue to be plagued by meaningless debates between the "New" and the "Old Look" or between "Formalists" and "Functionalists."

I have painted a gloomy picture of our ignorance, but I think the remedy is at hand if we will only grasp it. A research program that attempts systematically to test the limits of the several sources of variation in perceptual behavior should soon enable us to make intelligible distinctions between the role of need and the role of memory, between the role of sensory organization and the role of response set, between the role of sensation and the role of meaning. Such a program may even save us from concluding that sensory experience is more plastic than it really is.

REFERENCES

1. ASCH, S. E. Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership, and men.* Pittsburgh: Carnegie Press, 1951.
2. ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **6**, 183–193.

3. BRUNER, J. S., & GOODMAN, C. C. Value and need as organizing factors in perception. *J. abnorm. soc. Psychol.*, 1947, **42**, 33–44.
4. BRUNER, J. S., & POSTMAN, L. Perception, cognition, and behavior. *J. Pers.*, 1949, **18**, 14–31.
5. BRUNSWIK, E. Distal focussing of perception: size constancy in a representative sample of situations. *Psychol. Monogr.*, 1944, **56**, No. 1 (Whole No. 254).
6. BRUNSWIK, E. Remarks on functionalism in perception. *J. Pers.*, 1949, **18**, 56–65.
7. CARTER, L., & SCHOOLER, E. Value, need, and other factors in perception. *Psychol. Rev.*, 1949, **56**, 200–208.
8. ELLSON, D. G. Hallucinations produced by sensory conditioning. *J. exp. Psychol.*, 1941, **28**, 1–20.
9. GIBSON, J. J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.
10. GILINSKY, A. S. Perceived size and distance in visual space. *Psychol. Rev.*, 1951, **58**, 460–482.
11. HASTORF, A. H. The influence of suggestion on the relationship between stimulus size and perceived distance. *J. Psychol.*, 1950, **29**, 195–217.
12. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
13. HELSON, H. Adaptation-level as a frame of reference for prediction of psychophysical data. *Amer. J. Psychol.*, 1947, **60**, 1–29.
14. HESS, E. H. Development of the chick's responses to light and shade cues of depth. *J. comp. physiol. Psychol.*, 1950, **43**, 112–122.
15. HESS, E. H. Maturation and learning in the development of pecking accuracy in chicks. *Amer. Psychologist*, 1953, **8**, 367.
16. HOWIE, D. Perceptual defense. *Psychol. Rev.*, 1952, **59**, 308–315.
17. KLEIN, G. S., SCHLESINGER, H. J., & MEISTER, D. E. The effect of personal values on perception: an experimentive critique. *Psychol. Rev.*, 1951, **58**, 46–112.
18. KÖHLER, W., & WALLACH, H. Figural after-effects. *Proc. Amer. Phil. Soc.*, 1944, **88**, 269–357.
19. KOHLER, I. *Über Aufbau und Wandlungen der Wahrnehmungswelt*. Vienna: Rudolph M. Rohrer, 1951.
20. LAZARUS, R. S., YOUSEM, H., & ARENSBERG, D. Hunger and perception. *J. Pers.*, 1953, **21**, 312–328.
21. LEVINE, R., CHEIN, I., & MURPHY, G. The relation of the intensity of a need to the amount of perceptual distortion: a preliminary report. *J. Psychol.*, 1942, **13**, 283–293.
22. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
23. MILLER, G. A. The magical number seven, plus-or-minus two: some limits on our capacity for processing information. *Psychol. Rev.*, in press.
24. POSTMAN, L. The experimental analysis of motivational factors in perception. In *Current theory and research in motivation*. Lincoln: Univer. of Nebraska Press, 1953.
25. POSTMAN, L. On the problem of perceptual defense. *Psychol. Rev.*, 1953, **60**, 298–306.
26. PROSHANSKY, H., & MURPHY, G. The effects of reward and punishment on perception. *J. Psychol.*, 1942, **13**, 295–305.
27. RIESEN, A. H. Post-partum development of behavior. *Chicago Med. Sch. Quart.*, 1951, **13**, 17–24.
28. ROCK, I., & FLECK, F. S. A re-examination of the effect of monetary reward and punishment on figure-ground perception. *J. exp. Psychol.*, 1950, **40**, 766–776.
29. SCHAFER, R., & MURPHY, G. The role of autism in a visual figure-ground relationship. *J. exp. Psychol.*, 1943, **32**, 335–343.
30. WALLACH, H. Some considerations concerning the relation between perception and cognition. *J. Pers.*, 1949, **18**, 6–13.
31. WALLACH, H., O'CONNELL, D. W., & NEISSER, U. The memory effect of visual perception of three-dimensional form. *J. exp. Psychol.*, 1953, **45**, 360–368.
32. WALLACH, H., & AUSTIN, P. Recognition and the localization of visual traces. *Amer. J. Psychol.*, 1954, **67**, 338–340.

(Received November 25, 1954)

Psychological Review
Vol. 63, No. 1, 1956

EXTINCTION RE-EXAMINED AND RE-ANALYZED: A NEW THEORY

GREGORY RAZRAN

Queens College, New York

In 1939 the writer presented in this Journal (**42**) a 34-page analysis of the then-known facts and theories of extinction. The analysis has in the main —it seems—been accepted by Hilgard and Marquis (**17**), Spence (**50**), Woodworth and Schlosberg (**53**), and several other systematic writers in this area. Since 1939, however, a good number of new and basic facts on extinction have been uncovered to suggest a new—indeed a novel—theoretical alignment of the entire field. The facts—in 1955 even more than in 1939—continue to come predominantly from Russian laboratories. American animal laboratories of late have been, as is well known, chiefly occupied with operant conditioning which—without minimizing at all its systematic significance—is, certainly in its extinction phase, all too complex to allow unequivocal isolations of primary variables. On the other hand, theoretical discussions of extinction have been in recent years almost wholly confined to this country: Russian CR theorizing has been "frozen" for some time.

The purpose of the present article will thus be threefold. (*a*) It will present an up-to-date critical and succinct summary of the facts of extinction, basing the summary primarily (but by no means wholly) upon a statistical analysis of both old and new Russian classical CR experiments (97 in number). (*b*) It will show how present theories of extinction are, in the face of the reanalysis of the facts of extinction, even more inadequate than they were in 1939 (no new theory has really been added since). (*c*) It will offer a new theory of extinction which, it will argue, does account well for all the known facts of extinction. The presentation, discussion, and defense of the new theory (including specific suggestions for its further testability) provide the main purpose of the article.

Facts of Extinction: 1955

Space does not permit a detailed presentation and documentation of specific results of extinction under 31 headings, in the manner of the earlier article (many more headings would be needed now). Instead, the evidence will be compressed in seven summary statements which, it is hoped, will still retain all the essential findings in the entire area.

1. *General characteristics.* Extinction continues to be clearly a less than 100 per cent phenomenon. Instances of difficult and even impossible extinction are constantly reported by classical CR experimenters. The instances seem to fall primarily into two main varieties: (*a*) a low-level variety of "unextinction" in low aments (**27**, **39**), in decorticated animals (**7**, **37**), in animals low in the phyletic scale (**30**, **32**), and in reflexes low in neural organization (**9**); and (*b*) a high-level variety of extinction difficulties encountered in conditioning human *Ss*, particularly conditioning to verbal stimuli (**20**, **33**, **43**, **56**). Then, there is the wide gap between the extinguishabilities of alimentary and defense CR's in general.[1] The former are, other things

[1] The differences between alimentary and defense CR's must not be regarded as merely a matter of adience vs. abience. There are at least three other significant distinctions between the two: (*a*) In alimentary condition-

being equal, much more extinguishable than the latter (6, 34, 39), and defense CR's based upon strong US's such as intense electric shock may indeed be quite unextinguishable (48). And finally, three more significant findings should be mentioned. (*a*) De-oesophagized dogs appear to condition without extinction (7). (*b*) Repetition decrements in UR's, usually known as habituation or adaptation, are by all signs very slow in manifestation and are much more like the "unextinction" than the extinction of CR's (6, 16, 37). (*c*) In the very early stages of conditioning, the nonreinforcement decrement may not behave like typical extinction (no recovery, no generalization, and no "disinhibition"; 54, 55).

2. *Correlations between extinction and conditioning.* These correlations still may be said to be prevalently (but not invariably) negative only when they are based upon the extinguishabilities and conditionabilities of the same R in the same kind of animals. Dogs (54, 55), and to some extent also rats (22), and even fish (14) and crustacea (32), which are quick "conditioners," are as a rule slow "extinguishers" and those that are slow "conditioners" are quick "extinguishers" when the same R's are used. But the correlations are definitely much more often positive than negative when the conditionabilities and extinguishabilities of different R's in the same animals or of the same R in animals of different species are compared. Animals low in the phyletic scale are both slower "conditioners" and slower "extinguishers" than animals higher in the scale (40), low aments are slower in both variables than normal human *Ss* (39), and low-level reflexes are slower than higher-level ones (9). Here also two complicating factors must be taken in account: (*a*) the fact that alimentary CR's extinguish more quickly than defense CR's while differing but little in conditionability; (*b*) the evidence that in adult human *Ss* and in older children the correlations between conditioning and extinction are negative in some experiments (10, 31) and positive in others (11, 56), even when the same R is used in each experiment.

3. *Rate of extinction.* Besides the wide qualitative divergencies in extinction mentioned earlier, there is now no doubt that the rate of extinction is significantly related quantitatively to: (*a*) the duration of the intertrial intervals (54, 55, 56); (*b*) the strength (magnitude or latency) of the pre-extinction CR (34, 35, 55); (*c*) the closeness of the magnitude of the pre-extinction CR to that of the UR (55); (*d*) the age of the pre-extinction CR (17, 21, 54); (*e*) the complexity of the CS (41); (*f*) the semanticity of the CS (43); (*g*) the duration of the CS-US interval (delay or trace) of the original conditioning (17, 35, 54*f*, 54*o*); (*h*) the administration of sedatives (2, 34, 46); and (*i*) the administration of stimulants (*ibid*). The specific relationships appear by all signs to be: inverse with respect to variables *a*, *b*, *c*, *d*, *e*, and *f* (above); direct with respect to variables *g* and *h*; and inverse again with respect to variable *i*. Yet, with respect to variable *a*, there is one well-controlled experiment in dogs by Asratyan (4; one of the most careful workers in Pavlov's laboratory) with electric shock as the US in which extinction was found to be fastest with intertrial intervals of one to three min-

ing, withholding the US means depriving the organism not only of the energy of *S* but also of what might be called its mass (food, acid, or water ingested). (*b*) Alimentary R's are regularly and steadily exercised by organisms in daily nonlaboratory situations in the sense that defense R's, certainly particular defense R's, are not. (*c*) Alimentary R's are interactingly organized into a chained unitary and largely circular pattern, while defense R's are either disorganizedly diffuse or quite uninteractingly discrete and linear.

utes, slow with intervals over five minutes, and very slow with intervals of 10 to 20 seconds. And it may also be worth noting that (*a*) CR's the magnitudes of which approximate those of their UR's are practically impossible to extinguish (**55**), and that (*b*) generalized CR's extinguish much faster than conditioned CR's (**18, 35**).

4. *Temporal* (*spontaneous*) *recovery.* Plotting 47 instances (**54, 55, 56**) of temporal recovery of extinguished salivary CR's in dogs against nine different stages of extinction (keeping recovery time constant and using deciles as indices of extinction stages), the writer found (*a*) little, if any, recovery in the first or the first and second stages of extinction; (*b*) a maximum of recovery in the second or third stage; (*c*) a deceleration of the recovery afterwards; and (*d*) a near-zero but not a total absence of recovery in the last two stages of the extinction. (Hovland in this country found that a conditioned CSR in human *Ss* which had been extinguished to 65.1 per cent of its normal value *recovered* 12.6 per cent in 24 hours, while one extinguished to 90.5 per cent *lost* 7.8 per cent during the same time; 18.) [2] Again, both in the writer's own studies and in his treatment of data from Pavlov's laboratories, recovery beginning at maximum seems to go on for several days, thus pointing to its being more likely a forgetting than an inhibition process, that is, a forgetting of extinction rather than a dissipation of extinctive (or reactive) inhibition. True, the time dimensions in the temporal recovery of extinguished CR's are by far shorter than those involved in the forgetting of pre-extinction CR's. But it must be remembered that recovery conceived of as a forgetting of extinction can manifest itself only by virtue of its rate of forgetting being faster than the rate of forgetting of the pre-extinction conditioning (presumably because of Jost's law), and that this differential in rates would admittedly be much reduced and eventually nullified in the course of time.

5. *Extrastimulus recovery* (or disinhibition, a term that the writer does not favor because it labels a theoretical assumption as a behavioral fact). Like temporal recovery, this characteristic is far from being universal and is even more variable. Thus, when the pre-extinction CR is well developed and the extra stimulus is weak in intensity, the extrastimulus recovery is at maximum in the first stages of extinction and decelerates afterwards (**54, 55**). But when the pre-extinction CR is in its stage of development and the extra S is moderate or strong, the recovery is near zero in the first stages of extinction, accelerates afterwards, and then decelerates until it is near zero again in the last stages of extinction (*ibid.*). And, while there is reason to regard the first finding as the representative one, since the second case is complicated by decrements in the pre-extinction CR, there is also the fact that some extrastimulus recovery occurs also when the extra stimulus is applied, not jointly with the extinguishing stimulus, but between extinction trials.[3] In general, an extra stimulus administered during extinction might well be conceived of as producing four separate effects: (*a*) it interferes with the process of extinction, (*b*) it interferes with the still present conditioning, (*c*) it raises the general

[2] The two CR's in Hovland's experiment were not of the same type, one being a generalization CR; yet his findings are well in line with the writer's analysis of the studies from Pavlov's laboratory.

[3] When the stimulus applied between extinction trials is the US of the CR, the recovery of the extinction is particularly high, the phenomenon resembling the "CR sensitization" of Beritoff or the "pseudoconditioning" of Grether and of Harlow in original conditioning.

activity level of the organism, and (*d*) it restores feed-back CS's that have been lost during the extinction.

6. *Stimulus generalization.* When tested in later stages of extinction, stimulus generalization of extinction is quite similar to stimulus generalization of conditioning (**54** *b, g, j, k, n;* **55**). However when tested in earlier stages, this generalization differs from the latter in four important respects (*ibid.*). (*a*) The amount of the generalization is considerably greater in the case of extinction. (*b*) The gradient of the generalization of extinction contains fewer steps than the one of conditioning, indeed, in most cases there seems to be no gradient at all, all related CR's receiving about the same amount of transfer of extinction. (*c*) Stimulus generalization of extinction does not extend to "heterogeneous" CR's; that is, to CR's the UR's of which differ from the UR of the extinguished CR. (*d*) The stimulus generalization of extinction seems to be much more affected by drugs—sedatives restricting it and stimulants expanding it—than is the stimulus generalization of conditioning. Again, it should be pointed out here that, besides the generalization of extinction, non-extinguished CR's become further reduced when their CR's are applied simultaneously or near-simultaneously with the CS's of the extinguished CR's (**34, 35**). The degree of this reduction is a simple growth function of the degree of the extinction.

7. *Extinction stages.* The *early stage* of extinction is invariably an immediate and automatic phenomenon. *It requires no learning or practice.* Organisms that take dozens or scores or even hundreds of trials to form a CR begin extinguishing it right after the first nonreinforced trial. Cases in which extinction appears only after some trials are exceptional and can be accounted for by other processes, while reports of curves of immediate extinction appear by the hundreds. There can be no mistake about it! On the other hand, there is also little doubt that *full-fledged* extinction involves some *new* learning or *counter learning.* A stimulus and/or a reaction of an extinguished CR does not behave (unless the CR has wholly faded away —been forgotten—in time) like one that has never been conditioned. Instead, it clearly points to the coexistence of two kinds of learnings, an old and a new, running counter to each other. If the CR is only partly extinguished, the co-presence of the old and the new learning are overtly demonstrable. But even a wholly extinguished CR bears, by all signs, within itself a two-way CR connection.

INADEQUACIES OF CURRENT THEORIES OF EXTINCTION

As is well know, there are at present current in American CR literature three main types of theories of extinction: (*a*) inhibition (or fatigue), (*b*) counter conditioning (or interference), and (*c*) counter perception (or counter sign-gestalt-expectation). Counter-conditioning theories may further be divided into autochthonous ones, in which the counter conditioning involves only the response conditioned, and "heterochthonous" ones, in which the counter conditioning involves responses other than the one conditioned. The inadequacies of each of the three types of theories will be considered separately.

1. *Theories of internal inhibition and of reactive inhibition (or fatigue).* These theories have been criticized more often for their general conceptual framework than for their specific discordance with empirical data. Howell (**14**), Lashley (**29**), Beritoff (**6**), Konorski (**26**), and Anokhin (**3**) have long classed "internal inhibition" as very bad neurophysiology, and there is no doubt that a large number of psychologists, including

Lashley, consider "reactive inhibition" bad psychology. Elsewhere in this Journal (44) the writer characterized Pavlovian and Hullian inhibition as a Hegelian dialectic. And in a general way that is what it is: an assumption that an event by its very occurrence generates its opposite—bears, so to speak, within itself the very seed, the growing seed, of its own destruction. At a philosophical level one may of course be a follower of Hegel (even though by now his philosophical adversaries may well outrank his advocates). But when set up as a cornerstone of organismic reactivity in an experimental science, this "transformation into opposites" is obviously all too general, too ubiquitous, and too tractable and arbitrary, halving observables and doubling inferables, and tending to lead to considerable ambiguity and many paradoxes. Some of these paradoxes have been recently pointed out in this Journal (13).

An inhibition theory of extinction is, moreover, at variance with a good number of specific empirical findings, and fails to account for a number of others. Only the most significant empirical contradictions and failures will be mentioned here: (*a*) the absence of temporal recovery in the first stages of extinction; (*b*) the development of extinction with long intertrial intervals, 24 hours or more, when inhibition should supposedly have been dissipated; (*c*) the continuation of temporal recovery of extinction for a number of days; and (*d*) the many-mentioned instances in which extinction is difficult or impossible. A fatigue-like conception of inhibition is further weakened by: (*a*) the formation of full-blown extinction in very light and practically "workless" reactions, and the lack of difference between their extinction and that of "work" reactions; (*b*) the invariably negative correlation in classical conditioning between extinction and intensities of conditional stimuli and magnitudes of conditional responses; and (*c*) the finding that physical work actually restores extinction (1). Moreover, inhibition theories of extinction fail to account for (*a*) the wide differences between the extinguishabilities of alimentary and defense reflexes, (*b*) the absence of extinction generalization when the CR's are heterogeneous, and (*c*) the faster extinction (other things being equal) of CR's in animals higher in the phyletic scale. The last finding usually leads the Russians to view inhibition as a sort of "higher mental" process, a view which, besides its arbitrariness, would seem contrary to the fatigue-like conception of the process. The fact of course is that Pavlov long realized the empirical shortcomings of inhibition and supplemented it by another effect-reversing dialectic, namely, "induction": "positive induction" when the reverse effect occurs in an event adjacent to the inhibited event, and "auto-induction" when the inhibited event is itself reversed. Hull has not gone so far (except for the extra dialectic of the "negative habit, i.e., a habit of *not* doing something" [21, p. 282; cf. 13]); and it is indeed the writer's impression that a number of Hullians would not mind giving up inhibition for something better, more specific and testable, although a few may be too committed to change.

2. *Theories of counter conditioning or interference.* The most cogent general objection to these theories is contained in the last paragraph of the preceding section which may be restated as: Conditioning typically requires a number of trials to show *any* evidence of formation, while extinction begins immediately and could not thus conceivably be merely a function of *new* learning or conditioning. A dog that has taken, let us say, 27 reinforcements to secrete one drop of saliva in response to a bell could not possibly lose 18 drops of this

response through a mere one-reinforcement-old new or counter connection. New connections are not established so fast! Laboratory studies of counter conditioning quite often reveal the persistence of the earlier CR's (**54**, **55**), at least in the earlier stages of retraining. Moreover, there is also the question asked by the writer in his earlier article: Just what is the specific factor that starts behaviorally the counter conditioning or interference? Again, as might have been noted, a good number of the facts of extinction presented in the preceding section are hard to reconcile with a counter-conditioning theory. Among these facts we may number: (*a*) the many negative correlations between rates of extinction and rates of conditioning, (*b*) the negative correlations between rates of extinction and lengths of intertrial intervals, (*c*) the many cases of "unextinction," (*d*) the absence of generalization to heterogeneous CR's, and (*e*) the absence of temporal recovery (presumably forgetting of extinction in this theory) in the early stages of extinction. More could be cited.

3. *The theory of counter perception or counter sign-gestalt-expectation.* A good argument could of course be made for the view that organisms which perceive the stimulus relations in a CR situation are thereby capable of accelerating or retarding and perhaps even of initiating or nullifying extinction or conditioning. Moreover, extinction is by its very nature more conducive to the emergence of a new perception than is conditioning, since in extinction an event does not merely gain a new significance but the new significance is no longer what it used to be, affording cognitive comparisons and "aha" experiences. Nonetheless, as an over-all theory of extinction, counter perception appears to be quite untenable for at least three reasons. (*a*) Counter perception, like counter conditioning, needs time to be developed while extinction is an immediate process. In the writer's experiment on partial reinforcement of salivary CR's in human *Ss*, extinction became evident right at the start while hypotheses of "no more food" were not reported by the *Ss* until after five or six nonreinforced trials (**45**). And one would not expect dogs and rats and pigeons to form hypotheses more quickly. (*b*) Counter perception cannot obviously apply to extinction in decorticated animals, in crustacea, in viscerovisceral reflexes (slow as the extinction might be in these instances), and in "misinformed" human *Ss*. (*c*) Counter perception does not account well for the recovery and generalization of extinction.

A NEW THEORY OF EXTINCTION

The writer's new theory of extinction involves three main propositions, two corollary propositions, and three supplementary propositions. Of the three main propositions, the second is the most novel and crucial one, while the remaining two are in reality just revisions of earlier views. The acceptance of the first main proposition is independent of the acceptance of the latter two, each of which is likewise independent of the other. Again, the corollary propositions refer to incomplete extinction-like repetition decrements, while the supplementary propositions refer to mechanisms that accelerate extinction in *some* CR situations. (The mechanisms postulated in the supplementary propositions are held by some writers to account for all extinction—erroneously, according to the writer—and logically the first supplementary proposition could be substituted for the second main proposition and either the second or the third supplementary proposition for the third main proposition.)

The Main Propositions

Main Proposition I. Extinction is a successive two-phase process consisting of (a) an early stage of automatic deconditioning or unlearning and of (b) a later stage of counter conditioning or new learning.

A *successive* two-phase view of extinction clearly inheres in the general consideration that (if inhibition theories are rejected) the special characteristics of extinction stimuli which distinguish them from nonconditioned stimuli must obviously be acquired through learning which *ipso facto* presupposes a prelearning period. Moreover, such a view is empirically evident in the fact that the early and the late stages of extinction differ from each other in qualitatively distinct functional relationships. There is, for instance, little or no temporal recovery of extinction in its early stages and little or no extrastimulus recovery of the process in the later stages. And—what is more striking—in the later stages of extinction its generalization is quite like the stimulus generalization of conditioning, while in the early stages it is quite unlike the latter. Other evidence for, and advantages of, the present proposition will be brought out later; but it might be noted here that no bona fide *successive* two-phase view of extinction has ever been proposed. Hull's Postulate 9 refers, as is known, only to a summation (simultaneous or near-simultaneous) of reactive inhibition and conditioned inhibition, and Guthrie's concern has been with the initial reversal of cues in extinction rather than with an initial characteristic continuing phase of it. And, while the writer in his earlier article assumed that "extinction response decrement" and "extinction S-R decrement" ". . . are prerequisites and must therefore temporally precede all other factors in extinction. . ." (**42**, p. 285), he, too, had no clear view of the nature of the preceding decrement. An "automatic deconditioning" assumption could not be offered without some hypothesis about the mechanism of the deconditioning, which is the task of the coming proposition.

Main Proposition II. The automatic deconditioning in the early stage of extinction is a direct result of the loss of the interoceptive and the proprioceptive conditional stimuli (feed-back CS's) which in the original conditioning were an integral part of the CR situation and which when the unconditional stimulus is withheld and the evoked reaction is reduced cease to be present.

The Russians have clearly shown in the last 15 years that typical classical conditioning—alimentary, vasomotor, respiratory—is a function of a continuous interplay of exteroceptive and interoceptive-proprioceptive (hence to be known as interoproprioceptive) factors (**9**, **23**, **25**, **28**, **36**, **38**). Direct stimulation of any viscus (often executed through surgery) by any adequate stimulus could become a CS for the evocation, cessation, or reduction of any visceral or skeletal R in conventional conditioning manner (*ibid.*). The first stage of extinction may thus be conceived of as merely a case of conditioning to a compound stimulus in which some of the components of the compound, here the interoproprioceptive components, gradually drop out and thus gradually reduce the CR. The dropping out of the interoproprioceptive components, it may be necessary to add, is simply a function of the fact that the CR is smaller in magnitude than the UR. If the magnitudes of the two are close to each other, the extinction of the CR should be slow or impossible, which is exactly what known findings indicate (*supra*). Moreover, up to a certain limit, the reduction of the CR should proceed as a sort of self-

propelling mechanism: the reduced CR reducing the CS components and the reduced components reducing further the CR.

The present proposition fits so well many facts of the extinction of classical CR's with which the Russians are most familiar that one can attribute their overlooking it only to a reluctance—indeed, an inhibition—to depart in any way from Pavlovian orthodoxy. However, the facts accordant with the proposition are indeed many and impressive. Thus, it is easy to see why, according to the proposition, de-oesophagized dogs would fail to extinguish their CR's and why extinction would be slow and difficult in animals low in the phyletic scale and in low-level reflexes. Interoproprioceptive mechanisms would presumably be little, if at all, operative in all these instances. Likewise, the faster extinguishability of alimentary CR's as compared with that of defense CR's could be attributed to both the greater involvement of feedback CS's and the wider magnitude discrepancies between CR's and UR's in the former case. An interoproprioceptive concept of extinction would also predict fast extinction for delayed and trace CR's, and negative correlations between extinction and conditioning of the same R in the same species of animals. For delayed and trace CR's presumably tend to involve a wide matrix of interoception and proprioception, while animals that condition slowly would similarly involve more of their internal system in the conditioning and thus extinguish their CR's quickly, and those that condition quickly would involve less of it and thus extinguish their CR's slowly.

Of special significance to the plausibility of the proposition is the evidence of the generalization of extinction and the effects of extra stimuli and drugs. And here again the proposition would seem to accredit itself well. Both extra stimuli and stimulants (drugs), which no doubt increase interoproprioception and restore lost CS's, decrease extinction, while sedatives which decrease interoproprioception increase extinction. Some time ago Culler (12) argued that CS's could become virtually independent of their particular US's if only other extra stimuli supply the CR's with adequate motivation. However, experimental evidence—indeed, Culler's own data—shows only that extra stimuli may retard considerably the extinction of CR's. Moreover, the interpretation of this retardation need not invoke such a general concept as motivation or remotivation when a specific and testable hypothesis of reactivation of lost feedback CS's is available. Again, the special characteristics of the generalization of extinction in the early stages of extinction—namely, the larger amounts of transfer and the near absence of a gradient—would also seem to go well with the hypothesis of the present proposition, as one might rightly assume that the interoproprioceptive aspects of CR's have more "common elements" than do the exteroceptive ones. Then, there is the important finding that in its early state generalization of extinction does not extend to CR's that have different UR's, which could be explained only as due to an absence of common interoproprioception (4).[4] On the other hand,

[4] The interoceptive and proprioceptive CS's under discussion here refer to stimuli arising in all the activated effectors of a CR situation and not merely in those of the specific response conditioned (or measured)—e.g., in the effectors of the total act of eating and not merely in those of salivary action. Or, to put it somewhat differently, typical classical conditioning involves two kinds of CS's: primary, experimenter-manipulated external CS's; and secondary, response-initiated (or organism-supplied) internal CS's. And it is the secondary, feedback and response-bound CS's—or rather their loss—that are postulated here to be at the core of the automatic deconditioning stage of extinction.

the present proposition does not account for the temporal recovery of extinction. But, as it will be remembered, no such recovery seems to occur in the early stages of the process.

Main Proposition III. The later stage of extinction consists primarily of counter conditioning in which the to-be-conditioned response is the reduction (not the cessation) of the CR, the to-be-conditioned stimuli are chiefly the interoproprioceptive stimuli involved in the CR, and the entire process is often aided by intertrial conditioning in which interoproprioceptive and exteroceptive stimuli of preceding trials become conditioned to the response reductions in the succeeding trials.

The present conception of autochthonous counter conditioning differs from the related concepts of Guthrie (15) and of Hull (21) in three significant respects. (*a*) The counter conditioning is assumed to be not to a "no response" or to a "cessation of a response" or to "what not to do" but to a reduction of a response. The fact, of course, is that, as far as the organism and not the theorist is concerned, there is no "no response" or a "cessation of response" when the US is first withheld in the CR situation. There is only a reduction of a response, and by all logical and empirical signs such a reduction will become a to-be-conditioned datum. (*b*) The chief and prior (but not the only) to-be-conditioned stimuli in the counter conditioning are held to be the interoproprioceptive stimuli involved in the conditioning. The reason is simple. It is these stimuli that are first on the spot and are contiguous with the reduction of the CR. While in typical original conditioning the exteroceptive stimuli seem to be the first signaling agents and the interoproprioceptive stimuli come later, in extinction the sequence is assumed to be reversed. Conditioning places the organism, as it were, in commerce with the external environment, while extinction places it in commerce with itself. (*c*) Intertrial conditioning is postulated. In most extinction experiments, the duration of the intertrial intervals are within the limits of effective CS-US delays, so that a stimulus of a preceding trial may well become the CS of the R in a succeeding trial.

The concept of *autochthonous counter conditioning* was added to the concept of *interoproprioceptive automatic deconditioning* in the writer's theory of extinction on the basis of five specific considerations. (*a*) The experimental arrangement of extinction consisting of a progressive reduction of a response upon repeated applications of a stimulus at fairly short intervals (typically) would seem to be conducive, in accordance with our knowledge of conditioning, to the development of such a counter conditioning either in the same extinction trials or in intertrial successions. (*b*) The fact that extinction CS's reduce CR's other than their own indicates that these CS's have acquired in the course of extinction some definite counter-conditioning tendency, and it would seem to be better scientific methodology to ascribe this counter tendency to a fully testable concept of learning than to a "half-testable" one of internal or reactive inhibition. (*c*) Interoproprioceptive automatic deconditioning could account only for partial and not for total extinction, that is, only for the CR decrement resulting from the loss of the interoproprioceptive CS's in the S-R situation. (*d*) Automatic deconditioning does not account for temporal (spontaneous) recovery, the parameter of which would seem to suggest that it is a function of a differential in the rates of forgetting of two counter habits. (*e*) The functional similarity between stimulus generalization of extinction and of conditioning when extinction is in its later stage suggests a similarity of

mechanisms at this stage, i.e., that extinction is then a form of conditioning.

Moreover, while the writer's theory assumes the counter-conditioning phase of the extinction to begin after the automatic deconditioning phase has been in operation, and the automatic deconditioning phase to end well before the extinction is completed, there is, nonetheless, every reason to postulate a period in which the two phases operate simultaneously. Hence, a number of characteristics of extinction are likely to be a function of the mechanism of one phase or of the other, or of the two combined. Thus, the generalization of extinction may be thought of as due to either (*a*) the dropping out of common interopriprioceptive CS's in the automatic deconditioning, or (*b*) the operation of common interopriprioceptive CS's in the counter conditioning, or (*c*) both. Similarly, extrastimulus recovery of extinction may be a result of (*a*) the restoration of proprioceptive CS's in the automatic deconditioning, or of (*b*) the reduction of the counter conditioning through interference of distraction, or of (*c*) both. Again, while automatic deconditioning should produce faster extinction with short intertrial intervals, counter conditioning should be more readily effected through longer intervals, suggesting that extinction would be fastest if the intertrial intervals were short in the first stages of extinction and longer in its later stage. No such experiment has ever been tried, but there is no reason why it should not be, and there is little doubt that others like it could be devised not only to separate the effects of each of the assumed two extinction mechanisms but also to determine exactly their relative roles and magnitudes. In general, extinction illustrates well both the richness of variables in CR situations and the need for exact theoretical formulations to solve them.

The Corollary Propositions

Corollary Proposition 1. An extinction-like repetition decrement of a responce may be produced without interoproprioceptive deconditioning. Unlike extinction, however, such a decrement is very slow in manifestation, and there is at present no adequate specifically testable hypothesis to account for the occurrence of its initial stage.

This corollary proposition aims to cover two separate phenomena: (*a*) the slow extinction in lower animals, low-level reflexes, and CR's close in magnitude to UR's; and (*b*) the habituation or adaptation effects of UR's which, too, are very slow in manifestation. No specific experimental delimitation of the functional characteristics of slow extinction has ever been made. But habituation or adaptation have been known to manifest recovery and generalization (**16, 37**) which might be taken as evidence of some sort of autochthonous counter conditioning. There would seem to be no difficulty in asuming that habituation or adaptation involves some counter conditioning, but the possible nature of the mechanism that initiates the habituation or adaptation decrement remains a problem. A fatigue hypothesis without evidence of chemical changes is too general and inadequate, while the writer's early "asynchronization" theory is too neural and not of immediate value to behavioral theory. At any rate it should be stressed that earlier identifications of habituation-adaptation with extinction are not justified: the vast difference in rate of decrement makes the two quite distinct from each other.

Corollary Proposition 2. An extinction-like nonreinforcement which manifests none of the other extinction characteristics occurs in the very early stages of conditioning.

This phenomenon of a nonreinforce-

ment decrement without evidence of recovery and generalization has been noted by the writer, comparing results of nonreinforcements in a large number of classical CR experiments (**54, 55**). It appears to be a genuine phenomenon but one that is present only in the very early stages of conditioning, presumably before any degree of CR irreversibility has been reached. Its mechanism cannot be guessed at and perhaps need not concern us much as the phenomenon is obviously only passing in nature.

The Supplementary Propositions

Supplementary Proposition 1. In some CR situations, nonreinforcement may further reduce a CR when an R other than the UR, which during conditioning was held in abeyance by the UR, is released through the withholding of the US, and thereupon comes to interfere with the CR.

This is one of the two principles implied in interpreting extinction in terms of dominance and competing reaction systems (**8, 52, 42**). As a supplementary mechanism it is no doubt operative in a good number of CR situations. But as a universal account of extinction it is by no means justified. There are scores of bona fide extinction experiments in which careful observation and recording reveal no trace of the emergence of any competing reactions, and the general theory of paralleling "competing behavior systems" with "reciprocal innervation at the spinal level" is, to say the least, far-fetched and not readily testable. Moreover, the present proposition alone (even if combined with the one succeeding) does not do justice to a number of previously mentioned empirical facts.

Supplementary Proposition 2. In some CR situations, nonreinforced CS's become conditioned in the course of extinction to R's other than the UR. These other R's are either R's that have been held in abeyance by the UR during the conditioning or other R's that happen to be present during the development of the extinction.

This is the second principle implied in "heterochthonous" interference theories of extinction. And what has been said about the first principle of these theories applies equally well to the second: either principle is good as a supplementary hypothesis but poor as a universal rule.

Supplementary Proposition 3. In some CR situations extinction is greatly accelerated when the subjects perceive the changed significance of the CS when the US is withheld.

Like the first two propositions in this section, this one, too, is regarded by the writer as only supplementary. The reasons for the limitation of this proposition were given in an earlier section (*supra*, p. 44).

FINAL WORD

1. *The testability (verifiability) of the new theory.* The interoceptive and proprioceptive stimuli held by the theory to be the basis of the deconditioning and the counter conditioning in extinction are by no means postulated as any kind of intervening variables. They are observables! As already mentioned, in the last 15 years the Russians have clearly demonstrated both the existence of interoceptive and proprioceptive conditioning (particularly the former) and their interaction with typical exteroceptive conditioning. And while, as far as the writer is aware, no experiment has ever been specifically devoted to a comparison of the CR states of the viscera during alimentary conditioning and during alimentary extinction of conditioning, there is no reason why such an experiment could not be undertaken. Moreover, one could of course compare extinction of, let us say, alimentary conditioning in normal animals and in ani-

mals with de-afferented viscera. The isolation of proprioceptive stimulation presents more of a problem, no doubt, but could probably also be mastered.[5]

2. *Extinction of operant conditioning.* It has always been the opinion of the writer that a thorough analysis of primary variables in operant conditioning is difficult, if not impossible, without a prior analysis of such variables in classical conditioning. And it is only such a "prior analysis" which this article offers with respect to extinction of operant conditioning, even though a number of facts presented here have been directly verified, as is known, in operant CR experiments, and the theory, too, is not alien to operant CR thinking. More than that, the writer is by no means ready to accord priority to operant over classical conditioning with respect to theoretical and methodological—and even practical fundamentals of learning. The contrary appears to him to be true, and he indeed deplores the decline of classical CR experimentation in this country. Certainly, such extreme views as expressed recently in this Journal by one author (**47**) that "conditioning [meaning classical conditioning] is an artifact" are indefensible—and patricidal. They could have been put forward only in total unawareness of the almost 500 well-controlled experiments in classical conditioning for most of which the author's alternate "reinforcement" is a very contorted and unnatural or impossible explanation. True, a large number of these classical CR experiments are, unfortunately, unknown to American readers. Yet they do exist.

[5] The present formulation is not related to Guthrie's concept of movement-produced stimuli which pertains to the proprioception of the *original responses* to the CS's and not to that of their conditioned (and accompanying) responses. But it is related in a way to some of the views of Solomon and Wynne (**48**) on the relative failure of extinction in intense shock conditioning (traumatic avoidance learning in their terminology). However, these authors assume this failure—"partial irreversibility" in their terms—to be in some way a direct function of the *mere involvement* of feedback CS's in their conditioning whereas, according to the writer, mere involvement of such CS's is, on the contrary, the very basis of the development of extinction, and failure of extinction occurs only because intense shock conditioning—as any other intense conditioning—produces CR's of high magnitude and consequent *retention* of secondary CS's upon the withholding of US's. The writer's emphasis in this area of "unextinction" is on an intense conditioning and not solely on intense shock—or other abient—conditioning. It is an emphasis which suggests, if one desires, principles of "appetitive CR conservation" and "infant CR conservation" and even "semantic CR conservation" along with "anxiety CR conservation"—to borrow the term "conservation" from Solomon and Wynne but to use it as a mere description for outcomes of unextinction. (The instrumental aspect of shock conditioning is a separate issue which need not be raised here.)

SUMMARY

1. A critical up-to-date summary of the facts of extinction, based upon a statistical analysis of more than a hundred classical CR experiments, has been presented.

2. All three main current theories of extinction—inhibition, counter conditioning, and counter perception—have been found inadequate to account for the facts of extinction.

3. A new theory holding extinction to be effected in two stages and by two processes has been advanced. The two stages and processes have been postulated as (*a*) an early stage of automatic deconditioning due to the loss of feedback interoceptive and proprioceptive CS's when the CR is reduced, and (*b*) a later stage of counter conditioning in which the CR reduction is conditioned primarily (but not wholly) to interoceptive and proprioceptive CS's, and is aided by intertrial reduction conditioning.

4. The facts of classical conditioning

were thoroughly analyzed and found to fit well the new theory.[6] Specific suggestions of its direct verifiability were offered.

5. The reduction of classical CR experiments in recent years in this country was deplored.

REFERENCES

1. ABDULASE, K. [The effects of physical fatigue upon individually-acquired or conditioned reflexes.] *Russ. Fiziol. Zh.*, 1927, **10** (1–2), 169–176.
2. ANICHKOV, S. V. [The pharmacology of conditioned reflexes.] *Fiziol. Zh. SSSR*, 1952, **38**, 3–12.
3. ANOKHIN, P. K. [Key problems in the study of the highest nervous activity.] In P. K. Anokhin (Ed.), *Problemy vysshey nervnoy deyatel'nosti* (Problems of the highest nervous activities). Moscow: Akad. Nauk, 1949, pp. 9–128.
4. ASRATYAN, E. A. [The physiology of motor-defensive conditioned reflexes.] *Trud. Fiziol. Lab. Pavlova*, 1941, **10**, 25–39.
5. BASCHMURIN, A. E., & MULBERG, B. M. [The problem of the localization of neural centers in the cerebral hemispheres.] *Trud. Leningr. Vet. Inst.*, 1928, **2**, 85–88.
6. BERITOV, I. S. [*Individually acquired activity of the central nervous system.*] Tbilisi: GIZ, 1932.
7. BOLDYREFF, W. N. *Two new laws of cerebral function.* Battle Creek, Mich.: Battle Creek Sanitarium, 1931.
8. BOROVSKI, V. M. An attempt at building a theory of conditioned reflexes on spinal reflexes. *J. gen. Psychol.*, 1929, **2**, 3–11.
9. BYKOV, K. M. [The cerebral cortex and the internal organs.] *Moscow: Medgiz*, 1947. [German translation. Berlin: *Volk und Gesundheit*, 1953.]
10. CAMPBELL, A. A. The interrelation of two measures of conditioning in man. *J. exp. Psychol.*, 1938, **22**, 225–243.
11. CHUCHMAREV, Z. I. *Subcortical psychophysiology.* Kharkiv: GIZ, 1928.
12. CULLER, E. Recent advances in some concepts of conditioning. *Psychol. Rev.*, 1938, **45**, 134–153.
13. GLEITMAN, H., NACHMIAS, J., & NEISSER, U. The S-R reinforcement theory of extinction. *Psychol. Rev.*, 1954, **61**, 23–33.
14. GUSELNIKOV, V. I. [Certain characteristics of conditioned reflex activity in fish.] *Fiziol. Zh. SSSR*, 1952, **38**, 612–618.
15. GUTHRIE, E. R. *The psychology of learning.* (2nd Ed.) New York: Harpers, 1952.
16. HARRIS, J. D. Habituatory response decrement in the intact organism. *Psychol. Bull.*, 1943, **40**, 385–422.
17. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning.* New York: Appleton-Century, 1940.
18. HOVLAND, C. I. The generalization of conditioned responses. III. Extinction, spontaneous recovery, and disinhibition of conditioned and generalized responses. *J. exp. Psychol.*, 1937, **21**, 47–62.
19. HOWELL, W. H. Inhibition. *Physiol. Rev.*, 1925, **5**, 161–181.
20. HUDGINS, C. V. Conditioning and voluntary control of the pupillary light reflex. *J. gen. Psychol.*, 1933, **8**, 3–52.
21. HULL, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
22. HUNTER, W. S. Conditioning and extinction in the rat. *Brit. J. Psychol.*, 1935, **26**, 135–148.
23. IRAPETYANETZ, E. S., & STEPANOVICH, E. V. [Interoceptive signalization and the sympathetic nervous system.] *Fiziol. Zh. SSSR*, 1949, **35**, 483–490.
24. IVANOV-SMOLENSKY, A. G. The irradiation of extinctive inhibition in the auditory analyzer of the dog. *Trud. Fiziol. Lab. Pavlova*, 1927, **1**, 229–256.
25. KONORSKI, J., & MILLER, S. [Conditioned reflexes of the motor analyzer.] *Trud. Fiziol. Lab. Pavlova*, 1936, **5**, 119–278.
26. KONORSKI, J. *Conditioned reflexes and neuron organization.* Cambridge: Cambridge Univer. Press, 1948.
27. KRASNOGORSKY, N. I. Bedingte und unbedingte Reflexe in Kindesalter und ihre Bedeutung für die Klinik. *Ergebn. inner. Med. Kinderhk.*, 1931, **39**, 613–730.

[6] It needs, however, to be pointed out that the writer's theory is based primarily upon experiments in classical salivary conditioning and that other classical CR experiments are brought in only to fill in certain gaps. Yet such a manner of theorizing could hardly be called selective or "miniature" in nature for the simple reason that the experimental evidence on extinction of salivary conditioning constitutes about three-quarters of all the evidence in the area, and is, moreover, in itself quite voluminous.

28. KURTZIN, I. T. [Materials to cortico-visceral pathology.] *Trud. Inst., Fiziol. Pavlova*, 1952, **1**, 454–492.
29. LASHLEY, K. S. Learning. III. Nervous mechanisms in learning. In C. Murchison (Ed.), *A handbook of experimental psychology*. Worcester: Clark Univer. Press, 1934, pp. 456–496.
30. LEUTSKY, K. M. [Conditioned reflexes in normal and in decorticated frogs.] *Fiziol. Zh. SSSR*, 1929, **12**, 235–249.
31. MATEER, F. *Child behavior, a critical and experimental study of young children by the method of conditioned reflexes*. Boston: Badger, 1918.
32. MIKHAILOFF, S. Expériences réflexologiques. *Bull. Inst. Oceanogr.*, 1920, No. 379, 1–8; 1921, No. 398, 1–11; 1922, No. 418, 1–12; 1923, No. 422, 1–16.
33. OSIPOVA, V. N. [Unextinguished associated reflexes in children.] *Vopr. izuch. vospit. lichn.*, 1927 (1–2), 33–46.
34. PAVLOV, I. P. [*Collected works.*] (2nd Ed.) Moscow: Akad. Nauk, 1951. 5 vols.
35. PAVLOV, I. P. [Pavlov's Wednesday seminars.] Moscow: Akad. Nauk, 1949. 3 vols.
36. POGREBKOVA, A. V. [Intero- and exteroceptive conditioned respiratory reflexes and their interactions.] *Trud. Inst. Fiziol. Pavlova*, 1952, **1**, 103–115.
37. POPOV, N. A. [The problem of the extinction of reflexes.] *Izv. Azerbaydz. Univer.*, 1923, **3**, 230–241.
38. PSHONIK, A. T. [The interrelations of exteroceptive and interoceptive conditioned reflexes.] In K. M. Bykov (Ed.), *Problemy kortiko-vitzeral'noy patologii* (Problems of cortico-visceral pathology). Moscow: Akad. Med. Nauk, 1949, pp. 255–269.
39. RAZRAN, G. Conditioned responses in children. *Arch. Psychol.*, 1933, **23**, No. 148.
40. RAZRAN, G. Conditioned responses: an experimental study and a theoretical analysis. *Arch. Psychol.*, 1935, **28**, No. 191.
41. RAZRAN, G. Studies in configural conditioning. VI. Comparative extinction and forgetting of pattern and of single-stimulus conditioning. *J. exp. Psychol.*, 1939, **24**, 432–438.
42. RAZRAN, G. The nature of the extinctive process. *Psychol. Rev.*, 1939, **46**, 264–297.
43. RAZRAN, G. The extinction of conditionings to verbal stimuli. *Amer. Psychologist*, 1950, **5**, 486.
44. RAZRAN, G. Conditioning and perception. *Psychol. Rev.*, 1955, **62**, 83–95.
45. RAZRAN, G. Partial reinforcement of salivary conditioning in adult human *Ss*. *J. exp. Psychol.*, in press.
46. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.
47. SMITH, K. Conditioning as an artifact. *Psychol. Rev.*, 1954, **61**, 217–225.
48. SOLOMON, R. L., & WYNNE, L. C. Traumatic avoidance learning: the principles of anxiety conservation and partial irreversibility. *Psychol. Rev.*, 1954, **61**, 353–385.
49. SOLOVEICHIK, D. I. [Conditions conducive to the restoration of extinguished conditioned reflexes.] *Trud. Fiziol. Lab. Pavlova*, 1940, **9**, 270–276.
50. SPENCE, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951, pp. 690–729.
51. USIYEVICH, M. A. [The activities of the cerebral cortex and the work of the internal organs.] *Trud. Fiziol. Lab. Pavlova*, 1941, **10**, 51–156.
52. WENDT, G. R. An interpretation of inhibition of conditioned reflexes as competition between reaction systems. *Psychol. Rev.*, 1936, **43**, 258–281.
53. WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. (Rev. Ed) New York: Holt, 1954.
54. *Theses from Pavlov's laboratory*, in particular those by: (*a*) Babkin, (*b*) Bylina, (*c*) Chebotareva, (*d*) Degtyareva, (*e*) Dobrovolsky, (*f*) Eliason, (*g*) Gorn, (*h*) Grossman, (*i*) Kasherininova, (*j*) Kogan, (*k*) Krasnogorsky, (*l*) Perelzweig, (*m*) Petrova, (*n*) Potekhin, (*o*) Zavadsky, (*p*) Zeliony.
55. *Trudy Fiziologicheskikh Laboratorii Akademika I. P. Pavlova*, 1926–1945. 12 vols.
56. *Trudy Laboratorii Fiziologii Vysshey Nervnoy Deyatel'nosti Rebyonka*. [Transactions of Ivanov-Smolensky's *CR* Child laboratory.] 1930–1940. 6 vols.

(Received February 18, 1955)

Psychological Review
Vol. 63, No. 1, 1956

ISSUES DISTINGUISHING IDIOGRAPHIC FROM NOMOTHETIC APPROACHES TO PERSONALITY THEORY [1]

JOHN L. FALK

University of Illinois

In personality theory, the question has arisen as to whether there are one or two kinds of prediction (**5, 17**). Those who hold that two types of prediction are possible claim that there is the "statistical" or "actuarial" kind, which seeks for "regularities and uniformities *characteristic of a whole class* of objects" (**1**); and the "individual" or "clinical" kind, which attempts to reveal the individual pattern of the unique personality. Reliance on the former alone characterizes the nomothetic approach, whereas insistence on the necessity of both is characteristic of the idiographic approach.

Allport (**1**) contends that the two approaches overlap and complement each other. For example, the process of medical diagnosis is idiographic but dependent upon nomothetic bacteriology and biochemistry for information concerning specific diseases. It isn't *what* the disease is that is so important but *how* one has it. The idiographic view, as applied by the clinician, maintains that this "how" of the matter cannot be estimated by "actuarial" means, since it is dependent upon a host of conditions specific to the individual considered, and to the situation in which he is found. That is, we cannot assign probability 0.3 to a particular person as a statement of the likelihood of his riding the trolley next Sunday. He will either ride it or not ride it, and a thorough study of the person-in-his-life-situation will predict his behavior next Sunday toward the trolley as a rider (riding probability = 1.0) or a nonrider (riding probability = 0.0), the probability of 0.3 being considered a meaningless, actuarial fiction when applied to the particular case. Now it is true that an individual will not take 0.3 of a trolley ride. If we wish for specific prediction, other variables must be taken into account which effectively assign the individual to a much narrower subclass. This will permit a more refined estimate of his behavior and bring the probability value closer to zero or 1. The idiographic view maintains that these "other variables" must not be merely further actuarial data which assign the individual to narrower and narrower subclasses (**5**), but that we must treat the person as a unique, understandable configuration of factors—literally in a class by himself. Nomothetic science, on the other hand, works on the assumption that individuals "have" laws according to lawful processes, which also implies that the laws have completely lawful interactions.

Dymond (**5**) states the case for idiographic clinical procedures as follows: ". . . though psychologists are now able to do considerable predicting in normative terms the clinician often feels he needs something more, and that he can do something better than make a probability statement when predicting a particular individual."

But according to the nomothetic view, the clinician does not make *a* probability statement; he in fact compounds many such statements. For example, when making predictions concerning an

[1] After this paper was written, Dr. Paul E. Meehl's book (**15**) appeared. For an interesting discussion of related material and review of empirical studies the reader should consult this book.

anxiety neurotic, he utilizes not only an estimate dealing with the general class of "anxiety neurotics," but also estimates from the classes sex, age, education, parental identification, etc., *under the condition of* anxiety neurosis. Nomothetic procedure does not require the clinician to make his best possible estimate by predicting at the mean of a single class. There is no methodological reason why the implicit compounding of conditional probabilities in the clinic cannot be made nomothetically explicit.

Theories are most amenable to application when dealing with explicit, general relations, but this is not to say that implicit relations among variables cannot be grasped in some fashion and fruitfully applied by the clinician, much as the early shipbuilders utilized Archimedes' principle long before it was given an explicit formulation. On the other hand, the "statistical" approach to personality does not force narrower and narrower subclasses with a view to ending up with a complex typology, but merely hopes to define a multitude of class intersections in a quantitative fashion with the result that more explicit, and hence shrewder, guesses will be possible.

THE "UNIQUE" EVENT AND SCIENCE

Lewin (12) maintained that Aristotle's conception of the lawful event requires that it be one which occurs frequently; events which happen but once are merely chance occurrences. Within such a scheme individual differences have no place, for they are nonrecurrent phenomena if we inspect an array of individuals. In sum, lawfulness is revealed by repetition; the single (and perhaps exceptional case) is fortuitous and not amenable to comprehension by laws. In place of this Aristotelian orientation, Lewin advocated that we cease "to abstract from the situation," and attempt "to hunt out those situations in which the determinative factors of the total dynamic structure are most clearly, distinctly, and purely to be discerned." Process dynamics are revealed by "the relation of the concrete individual to the concrete situation."

Although Lewin's opposition to Aristotelian conceptions is cogent, it is doubtful whether an appeal to fuller concretization, together with a turning away from abstraction, will produce the general laws which Lewin sought. Such directives are acceptable if the intent is merely to break through an existing set of tyrannical abstractions. But the ultimate goal is a better set of abstractions, not an actual return to the concrete situation. Current personality theory cannot predict certain aspects of behavior, but assumes that they are nonetheless lawful. By treating the issue as one between the lawful status of *recurrent* events on the one hand, and *unique* events on the other, Lewin has set the problem in an unfortunate way. The truly unique event (if there be such) cannot be treated within the framework of science. Events like novel mutations, supernovae, etc., can be so treated, since these are simply occasions within certain event classes whose temporal onset is at present not accurately predictable. The Aristotelian "fortuitous case" is shown to be lawful simply because it is *not* unique and can be relegated to a predictable class of occurrences. Recurrence does not imply the repetition of the same event (which is clearly an impossibility); it implies only that some select characteristics of the event can be noted on other occasions when similar events are taking place. It is, of course, the recognition of what constitutes a "similar event" that gives the scientist trouble. It took a Newton to recognize that the movements of falling bodies, the tides, and

the planets were similar events, and the explicit recognition was due to the forging of a powerful, new abstraction, not to the concreteness of a falling apple.

The requirement of Aristotelian science that repetition be a necessary condition in order that the event be included within the realm of the lawful was false insofar as it called for the recurrence of phenomenally similar events. If only one meteor had ever been observed, this would not constitute it a unique event for present-day science, since the class "occasions of combustion" has been formulated, and the conditions of class membership explicated. Modern science requires repetition not for the sake of ensuring that the event is not fortuitous (although this is a factor in situations calling for the application of statistical inference) but because repetition is the by-product of class formulation and repeatable verification. Repetition, then, is only an indirect requirement stemming from the necessity of abstracting a variety of events into one class, and from the potential replicability of any reliable procedure and result. Laws do not gain strength with repeated application, in the manner of a set of muscles. The "fortuitous," "unique" event of Aristotelian science has given way to the realization that such events exhibit lawful behavior when the correct abstractions are applied. The complexities of "unique" personalities await lawful unification by some psychological Newton.

WHOLES AS REDUCIBLE AND PREDICTABLE

London (13) has drawn an analogy from the fact that many properties of hydrocarbon molecules (e.g., heats of fusion, vaporization, hydrogenation, combustion, melting and boiling points, viscosity, density, color, refractive index, chemical affinities, and absorption spectra) are deducible from the properties of their constituent atoms. But a great number of these atomic properties are derived by inferring their modes of interaction in molecules. There is nothing in the credo of the reductionist which insists that parts be studied exclusively in isolation. The fact is that in their interactions the parts reveal certain of their properties which are not evident in isolation. An analysis into parts does not require that a subsequent synthesis of these parts be merely "and-summative" or simply additive. The rules of part combination make the deduction of whole-properties possible. As in chemistry, and in every discipline which seeks to deal with certain whole-properties, analysis and resynthesis are prerequisites to complete theory, not shifts to other "levels of analysis." The concept of an "electric current" may be operationally defined in terms of certain magnetic, thermal, or chemical effects. That these different defining operations do not constitute different concepts is shown by the fact that all these effects may be deduced from a unitary theory of electricity, magnetism, electrolysis, and heat (6). The properties of "electric current" as given by electrodynamics depend upon a set of operations and their theoretical interactions.

Thus, in the case of hydrocarbons, their properties may be deduced within the same framework by means of which the properties of the constituent atoms were ascertained. In the same fashion, the "electric current" is explained in terms of operations and their theoretical cross linkages, without resort to some more "molar level" for a completion of the story. Neither organic chemistry nor electrodynamics requires a change in explanatory "level" to account for phenomena having either a larger spatial scale [2] or more general effects.

[2] The reason why a scale model of a bridge may stand up quite nicely while breaking

Insofar as these physical analogies are germane, reductionist faith seems to have paid off rather handsomely. But even a cursory glance at the way in which reduction typically occurs is likely to prove disconcerting to the simon-pure "molecular level" theorist. Typically, the study of whole phenomena, produced by an aggregate of interacting "molecular" effects, gives rise to the formulation of general contexts which define the modes of interaction occurring at the reduction level. For example, the development of atomic theory would have been impossible without an extensive knowledge of chemical effects. Although textbook logic prefers to derive complex phenomena from a consideration of their simpler constituents, scientific practice very often makes its initial observations and inferences about the structure and function of simples (which are often unobservable) from a study of the complex. Thus, a study of molar phenomena throws light upon molecular processes by setting up certain systemic limits within which part synthesis must take place. Also, molar phenomena reveal the properties of "molecular" processes by exposing them to situations in which often unexpected transactional capacities of these parts are discovered. From a knowledge of these transactional possibilities of parts *the properties of entire classes of wholes are predictable.*

The purpose of these remarks has been to dispel the notion that the whole must be treated with a different brand of logic, or as a higher-level organization not amenable to reduction. Chemistry and physics do not stand in the relation of whole and part to one another. They are separate as to their techniques and objectives, not in terms of their factual data. In a similar vein, personality theorists who hold that the data of the "unique individual" are somehow different from the data of general psychology fail to realize that it is the operations performed, and not the type of data, which distinguish such areas of investigation. The intent of the personality theorist differs from that of the general psychologist (as does the chemist's from the physicist's), but his data are potentially predictable by the general psychologist, just as the chemist's are by the physicist. Let us hope that future personality theory will be as fruitful to psychology in the suggestion of critical data and constructs as the chemist has been to the physicist.

Before leaving this phase of the story, it may be well to incorporate the methodology of a science somewhat different from physics or chemistry. The idiographic proponent may feel that it is possible to prove the complete adequacy of a nomothetic approach to wholes for the physical sciences, but that personality is so dependent upon essentially nonrepeatable, historical processes that reduction is simply not feasible, each personality being a unique, emergent whole arising from the diverse chains of events composing its historic route. A science which closely parallels this state of affairs described by the idiographic viewpoint is archaeology. Consider this observation [8] by Carter and Mace (4, p. 179):

"There are numberless derelict objects in the storerooms of our museums which would give us valuable information could they but tell us whence they came, and box after box full of frag-

from its own weight when built full size can be completely explained in terms of a single fact covering this phenomenally striking difference: surface increases as the square, while volume increases as the cube, of the linear dimensions. No recourse to another "level" is called for.

[8] Compare with Freud (7): "A collection of dreams without associations and knowledge of the context in which it was dreamed does not tell me anything. . . ."

ments which a few notes at the time of finding would have rendered capable of reconstruction."

In this sort of situation the parts are present but can only be studied in isolation, the context being irrevocably lost through the carelessness of the digger or collector; i.e., the rules of combination are unfortunately past recovery. A comparable situation in physics would be one in which sets of observations taken under various conditions had been preserved, but the specifications for the conditions had been lost. In both cases all the results are present, but are not open to interpretation in the absence of information of those circumstances which gave rise to them. The only fundamental difference here is that the physicist, now knowing that certain outcomes are possible, might concentrate his efforts on certain experimental procedures and find out which condition produces which result, while the archaeologist cannot replicate. Neither case requires an interpretation in terms of "emergents" unrevealed by parts.

Where the archaeologist does have information of conditions and parts he is often able to predict quite well on the basis of past cases. When Carter and the Earl of Carnarvon discovered the tomb of Tut-ankh-Amen, they proceeded to photograph, catalogue, preserve, and clear the "wonderful things" in the antechamber, a procedure which took many seasons to accomplish. They had no doubts that they had uncovered the tomb of Tut-ankh-Amen, although they did not wish to penetrate into the actual burial chamber until the antechamber had been cleared. At this time others advanced various alternative hypotheses about the nature of the find, all claiming that it was not a tomb at all. To these objections Carter and Mace (4, p. xi) replied that ". . . so far as we have gone we have found nothing that should not belong to the funerary equipment of the king. All the objects are in perfect keeping with the evidence and knowledge gleaned from the fragmentary material of the royal tombs of the New Empire discovered in this Valley, and they are in every way pure late Eighteenth Dynasty Egyptian."

Later work on their part, subsequent to the publication of Volume I (4), bore out their contention. Their prediction of the whole-property (tomb of Tut-ankh-Amen) was based upon the properties of the classes of objects they had discovered, properties in turn derived from a study of other individual tombs in the Valley of the Kings.

The ideas of holism and emergence are intimately related. It is the alleged emergence of new wholes at higher levels of organization that is said to require "laws which are unique for each level" (16) to describe the situation adequately. Again, a consideration of which criterion of adequacy the investigator is utilizing is a long step toward the resolution of this question. If "adequacy" is meant in terms of completeness of explanation (rigorous reduction, detailed prediction), then laws entirely unique to "emergent levels" are meaningless; while if "adequacy" is construed as a pragmatic handling of certain molar phenomena, then reduction becomes unnecessary, and laws dealing with the operation of certain aggregates claim a primary place. The gas laws are reducible to thermodynamic principles, but such reduction adds little or nothing to the majority of situations where the gas laws are applied. For a more complete explanation (in the above sense) the principles of thermodynamics outdistance the gas laws by far.

In sum, "emergent wholes" do not give rise to sets of laws which are unrelated to the laws dealing with constituent elements. The wholes are constituted by the experimenter, and may

be treated apart from any reduction level, depending upon what classes of operations mesh with the purposes of the investigator. Formulation at some ultimate reduction level is neither the simplest nor most economic way to handle most problems, but this does not mean (as the history of science has shown) that we are dealing with unique, irreducible emergents. The way in which the simples are pictured as forming a whole serves to predict accurately complete classes of whole-properties. The holist has no difficulty in imagining the fivefold symmetry of a pentagonal crystal. The fact is, however, that crystal symmetry is built up from molecular structures, and it is impossible for a repeating pattern of these to be pentagonal (14). Thus, the nonexistence of one type of whole, i.e., pentagonal crystals, is as predictable from strictly molecular properties as is the occurrence of a multitude of other crystalline forms.

Misuse and limitations of reduction. It should be realized that the feasibility of reduction does not automatically validate the use of arguments by analogy. For example, although we may believe that personality is ultimately reducible to a set of physiological variables, the personality theorist is not thereby given carte blanche to utilize physiological constructs to explain away certain classes of behavior. Reduction processes are actually circumvented by attempting to apply constructs and equations, formulated within one realm of operational procedures, to phenomena arrived at by totally different modes of operation. To cite an extreme example, the fact that sociology is in principle reducible to biology or physics in no way validates the use of "social organism," "the principle of indeterminacy" (in sociology), "social inertia," etc. These may function as suggestive metaphors, but they do not lend a biological or physical stamp of approval to the user. By the same token, the hope that sociological data may somehow be read into existing physical equations is chimerical; reduction is not carried out by the blind application of the particular formal systems utilized by physics. Any set of laws is grounded to an operationally clear-cut domain, but this does not mean to say that these laws cannot be derived from a more basic set of axioms. However, the reduction of a set of laws to some more basic set of axioms seldom, if ever, means a simple identification of isomorphic structures, in which the x, y, z of one system conveniently becomes the x', y', z' of another. Yet certain analyses assume that this is the case: e.g., Spengler's treatment of civilization as a true organism, Haret's (8) *Mécanique Sociale* with its literal interpretations of classical mechanics—"center of gravity of a social body," "equality of action and reaction," "the conservation of energy of a social body," etc.

Hoagland's (9, 10) attempt to reduce the performance of temporal judgments to the action of some unidentified "chemical clock" enzyme by directly transforming "time perception" into chemical "rate of reaction" is allowable as a construct, but does not constitute reduction merely because both terms involve the concept of rate. What this implies for the development of personality theory is precisely this: For personality theory *qua* personality theory physiology is not causally prior to behavior. Much correlation and inference must proceed among phenomena within a single operational realm before reduction can become meaningful or even possible. It is fruitless for the psychologist passively to await "the solution" of problems such as schizophrenia and drug addiction by the biochemist. Even if we knew that a specific enzyme inhibitor would directly produce one type of schizophrenia, this

reduction would mean little for personality theory without considerable reliable knowledge of schizophrenic phenomena and their relations to other modes of behavior, cultural circumstances, etc. Thus, a theory of personality must stand or fall in terms of the validity of its operationally delimited constructs. Reduction to physiology cannot offer a more comprehensive explanation of personality unless the reduction process is subsequently reversed, and *predictions within personality research,* made in the light of reduction, are confirmed.

THE INTENTIONS AND CAPACITIES OF IDIOGRAPHIC AND NOMOTHETIC METHODS

It might be wise to stop and clear up a minor but insidious point. One is often subjected to the jibe that the psychologist can give the layman no guiding maxims in the evaluation of personality, and in fact shows himself to be rather mediocre in his evaluations of friends and associates. The implication is either that personality cannot be treated within the domain of science or that psychologists are fraudulent, misguided, etc., in that they do no better at evaluation than other people.

Allport (1, p. 5) states: "This exclusion of the individual from pure psychology has led to many anomalies. It has, for example, often been pointed out that the psychologist, in spite of his profession, is not a superior judge of people. He should be, but his ascetic and meager formulae derived from 'generalized mind' do not go far in accounting for the peculiar richness and uniqueness of minds that are organic and single."

The above argument is perhaps somewhat misleading. The market-place situation divorces the scientist, be he physicist, biologist, or psychologist, from his most important tool: the administration of standardized tests in controlled situations. Leaving this point to one side, we do not expect the organic chemist to be a judge of fine wines and cheeses merely because he knows something about the molecular structures which wring praises from a connoisseur. If we embark on a fishing trip in the Canadian north woods, we do not employ a geographer, geologist, or even an ichthyologist in order to meet the rigors of canoeing, camping and fishing, but we do hire a guide. The guide and the connoisseur have a more or less practical knowledge of a rather select set of particular whole-properties, while the scientist is concerned with somewhat different aims and methods. To expect the psychologist to be either a connoisseur of or a guide to personalities without (*a*) the benefit of employing the tools of his trade, or (*b*) the advantage of having any compulsory stake in the classes of discriminations which guides or connoisseurs apply, is an unreasonable request. If our hypothetical psychologist actually is a good immediate judge of people, it merely attests to his being a rather more complete human being.

When Windelband sought to divide science into the two disciplinary modes, nomothetic and idiographic, he had in mind not personality theory, but history. Historical explanation was not to be gained through the development of general laws, but was to concern itself with the understanding of some particular event in nature or society. Allport (1), while admitting both procedures as valid methods for understanding personality, views personalities as just such unique, particular events. The idiographic approach aims to present the individual in such a way as to preserve his "unique" qualities. A personality may be characterized in this fashion, but such qualities are interpretive constructs, not raw facts. Their admissibility into the scientific arena

cannot be questioned; their necessity in a complete theory of personality is, however, open to methodological doubts insofar as they are incapable of giving rise to testable theorems.

It is unfortunate that personality theory has become cloven along the individual-general axis. Neither method is entirely independent of the other, although the general theory approach is logically (though not historically) self-sufficient. We shall attempt to indicate that the methods have different objectives, but that nevertheless they are not actually independent.

The scientist interested in developing a general theory of personality is concerned primarily with arriving at sets of postulates with testable theorems, statements of the predictive accuracy of the theorems, and their quantitative expression. Allport (1) characterizes this approach as seeking "regularities and uniformities *characteristic of a whole class* of objects. . . ." But nomothetic science does not necessarily deal with characteristics which are common to all the members of a class. It may utilize the pure case (e.g., Newton's first law of motion) in which event *no* member can be observed to exhibit the characteristic (continued motion in a straight line). Individual differences in the motion of various objects are accounted for by other general factors (e.g., air resistance) in such a way that the "regularities and uniformities" are embodied not in the observed events themselves but in the theoretical structure. The realization of this fact removes at one stroke the alleged inadequacy of general theories in the handling of individual differences. Stated another way, general personality theory does not demand uniformity of response among subjects when dealing with some particular lawful relation. The relation may exist as an unobservable pure case, and of course the interaction of the relation with other lawful relations could easily serve to create great dissimilarity at the response end. Nomothetic science is adequately equipped to handle such interaction effects. The fact that hyperthyroidism may give rise to ophthalmic, metabolic, or cardiac disorders, and that in general ". . . the classical catalogue of discrete disease entities may be no more than a continuous spectrum of the possible outcomes in different individuals of the same disordered defense mechanism . . ." (11) has not turned endocrinology idiographic but only widened its nomothetic considerations.

The function of an idiographic approach is not to develop and test postulate systems, nor is it primarily concerned with the quantitative representation of relations; rather, it is the process of interpreting one set of complex responses in terms of another which claims the center of the stage. For example, the neurotic behavior of an individual may be gauged by interpreting his symbolic systems via dreams, word associations, etc. Here, the complex responses obtained in the clinical setting are of minimal interest per se. They are interpreted rather than measured, being treated as symbolic indicants of certain covert processes. The intent of the two approaches differs rather sharply, one being oriented toward precise measurement of variables, the other toward comprehensive interpretation [4] of particular cases (3). Clinical interpretation has developed in

[4] The theoretical structures erected by nomothetic science are also, of course, purely interpretive. The difference is that general theoretical interpretation must make predictions about constituted classes of events, while with idiographic interpretation of the single case the process often terminates with the formulation of the interpretive judgment. Although a specific idiographic prediction is confirmable (e.g., "This patient will commit suicide if event X occurs"), precisely *what* it is which is thus confirmed when X occurs and suicide results is not at all clear when we eliminate the omnipotence of the therapist as an alternative.

the relative absence of controlled study of the functional relationships among variables and their interactions—and yet it has come a long way. The question that we must now consider is a methodological one: is there of necessity a schism between idiographic and nomothetic methods? If not, do they simply complement one another, or is one reducible to the other?

It has already been indicated that idiographic methods are of considerable use in the location and rough estimation of new or crucial variables. In this way, such investigations are closely akin to pilot studies. Initial, trial experiments and the noted implications of laboratory happenstance are idiographic. It would appear that, insofar as the idiographic method is propaedeutic to scientific investigation, it is of prime value but not actually part of the methodology of science utilizing an explicitly formulated set of operations for predicting events. If the aim is the interpretation of one given complex, implicit event (personality) which continues to occur, then the behavior sampling might well consist in controlled testing, with the interpretation depending upon what is already known (viz., nomothetically) about the distribution of scores on the tests. The ultimate interpretation in terms of score configuration may be a leap to the literary, based upon implicit sources of information, but it is made out of necessity only as long as there is a lack of information concerning score interaction within various subclasses of subjects.

From these considerations we may conclude that there is no split between the methods as far as their status *as scientific methods* is concerned—both reduce to the nomothetic in this instance. While the idiographic contributions to scientific methodology may be reducible to the nomothetic background upon which they are viewed, it is, paradoxically, in its uncontrolled but novel observations of behavior that idiographic studies make their most valuable contributions to personality theory. Stated another way, we turn to idiographic procedures for the provision of working hypotheses, not for their confirmation. The answer, then, to our methodological question must be put as follows: there is no schism between the methods as they operate within scientific methodology, idiographic goals being ultimately and more precisely attained by nomothetic means. Outside methodology, the idiographic mode provides a fruitful supply of new variables and hypotheses, and is in this fashion complementary to nomothetic methodology.

The nomothetic viewpoint seeks to derive individual personality patterns as particular cases of general principles. It emphasizes the isolation of variables, since it is felt that this makes their accurate manipulation possible and that first things must come first, the simple analysis preceding the complex. There is, of course, no inherent reason why nomothetic procedures cannot handle interaction effects.[5] That a more comprehensive intercourse among variables in general personality theory has not occurred results more from maidenly inexperience than from any innate lack of potential excellence. The idiographic view stands for a resolute grasping of wider areas of the network of relationships, despite the lack of reliable information. Like venturing virgins, idiographers do not usually meet the mark in an exacting test of performance, but are capable of making interesting discoveries.

[5] Beck maintains that only idiography can fill the bill in this respect: "A universe of traits, variables in mutual interplay, affecting one another, these are the individual. This is the task which the idiographic method undertakes" (2).

SUMMARY

1. General laws and their interactions are potentially a sufficient structure to account for the "unique" personality.

2. Idiographic methods have a valid function in those situations where the worker wishes to take account implicitly of a large number of variables for which no interrelationships have been explicitly mapped. The fruitfulness of the method may be measured by (*a*) the validity of the clinical interpretations which it generates (i.e., the predicted event occurs, or somehow the patients improve), and (*b*) its discovery of new variables in behavior.

3. Inspection of a few examples from various scientific disciplines seems to indicate that whole-properties are potentially reducible and predictable. A study of whole-properties (i.e., phenomena as defined by operations considered to be somewhat more molar) reveals information concerning parts and their modes of interaction. With the aid of such considerations, generalizations about part-processes are formulated which serve to predict the properties of entire classes of wholes without the use of "emergents," "higher levels of analysis," and the like.

4. Laws handling the relations among molar aggregates are often of greater practical use than a consideration of the same phenomena in terms of more detailed reduction.

5. Reduction is of little use to the area of investigation undergoing reduction unless the reduction formulation is able to predict and bring within its explanatory compass other phenomena associated with the area being reduced.

6. Aside from the fact of accepting a unifactorial or bifactorial theory of prediction, respectively, the nomothetic and idiographic approaches differ in intent; the one aiming at the quantification of variables, the other striving to represent and interpret the pattern of single personalities.

7. Methodologically speaking, the nomothetic is potentially capable of usurping the realm of the idiographic. The chief value of idiography is seen as its function in the delineation of new variables and working hypotheses.

REFERENCES

1. ALLPORT, G. W. *Personality: a psychological interpretation.* New York: Henry Holt, 1937.
2. BECK, S. J. The science of personality: nomothetic or idiographic? *Psychol. Rev.,* 1953, **60,** 353–359.
3. BINDRA, D. The case study: its place in personality research. *Canad. J. Psychol.,* 1952, **6,** 11–17.
4. CARTER, H., & MACE, A. C. *The tomb of Tut-ankh-Amen.* Vol. I. New York: George H. Doran, 1923.
5. DYMOND, ROSALIND. Can clinicians predict individual behavior? *J. Pers.,* 1953, **22,** 151–161.
6. FEIGL, H. The mind-body problem in the development of logical empiricism. *Rév. Int. de Phil.,* 1950, **4,** 64–83.
7. FREUD, S. Letter to André Breton. Quoted by Nicolas Calas, in Surrealist intentions. *Trans/formation,* 1950, **1,** 48–52.
8. HARET, S. C. *Mécanique sociale.* Paris: Gauthier-Villars, 1910.
9. HOAGLAND, H. *Pacemakers in relation to aspects of behavior.* New York: Macmillan, 1935.
10. HOAGLAND, H. The chemistry of time. *Sci. Mon.,* 1943, **56,** 56–61.
11. LEVAY, D. Hans Selye and a unitary conception of disease. *Brit. J. Phil. Sci.,* 1952, **3,** 157–168.
12. LEWIN, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1935.
13. LONDON, I. D. Some consequences for history and psychology of Langmuir's concept of convergence and divergence of phenomena. *Psychol. Rev.,* 1946, **53,** 170–188.
14. LONSDALE, KATHLEEN. Determining atomic patterns. *Sci. News,* 1950, **18,** 59–76.
15. MEEHL, P. E. *Clinical versus statistical prediction.* Minneapolis: Univer. of Minnesota Press, 1954.
16. NOVIKOFF, A. B. Integrative levels in biology. *Science,* 1945, **101,** 209–215.
17. SARBIN, T. R. The logic of prediction in psychology. *Psychol. Rev.,* 1944, **51,** 210–228.

(Received March 9, 1955)

Psychological Review
Vol. 63, No. 1, 1956

A MICROMOLAR APPROACH TO BEHAVIOR THEORY

FRANK A. LOGAN

Yale University

A theory of learned behavior may be evaluated in terms of its success within the specific area for which it was designed; it may also be evaluated as a basis for integrating the several related behavior sciences. The first part of this paper is concerned with two questions: (*a*) Is the classical Hullian model (**3, 4, 5**) generally useful within its original area? and (*b*) Is this theory likely to be useful in other areas, such as economic behavior? In both cases, the questions raised involve fundamental features of the model and not that specific content which is readily modifiable on the basis of continuing research.

Since it seems to the writer that both of these questions must be answered partially in the negative, it may be profitable at this time to examine a potentially more useful model. Accordingly, the second and main part of this paper deals with what Hull called the micromolar approach to behavior theory. In a micromolar theory, quantitatively different behaviors (e.g., different amplitudes, speeds, etc.) are treated as different responses, rather than as different strengths of the same response. This somewhat different approach has not as yet been developed sufficiently to permit a complete statement of a formal micromolar model. However, some of the characteristics of such a theory can be stated and a few illustrative solutions can be made. These, it is felt, justify serious consideration of the micromolar approach.

The Macromolar Approach, as Exemplified by Hull

Most psychologists are familiar with the concept of molar behavior as distinguished from molecular behavior. Qualitatively different molecular responses are classified together in a single molar category; for example, the molar response "bar depression" includes a number of responses distinguishable at a more molecular level, such as depression with the right paw, left paw, etc. Molar behavior theories make predictions about the occurrence of *any* member of such a category, and are not concerned with possible distinctions among the members. The more molar analyses accordingly suppress a certain amount of information about the subject's behavior, but are typically employed either because a more molecular analysis is not readily available or because the information suppressed is simply not important for the purpose at hand. Whatever the reasons, social scientists typically deal with relatively molar classes of responses.

Once the decision concerning qualitative response classes is made, there still remain variations of a quantitative nature, such as the dimensions of amplitude, speed, etc. A common procedure, exemplified by Hull, is to consider variations in all of these dimensions as functions of a single factor, response tendency. This approach may be called the *macromolar* one.

Two general implications of a macromolar system concern us here, and these implications stand quite independently of the particular content of the model. First, the response dimensions (amplitude, speed, etc.) should improve monotonically with learning, other things constant.[1] That is, as learning trials are

[1] This statement is true only because the response dimensions do typically improve

performed, with other factors such as drive and reward invariant, each one of the dimensions of performance should improve until, perhaps, it eventually levels out asymptotically.

In the second place, stable and high rank-order correlations should be found between the several response dimensions. Since amplitude, speed, etc., each varies with a single factor, namely, response strength, a change in one must be accompanied by a corresponding change in the others.

Neither of these implications appears to be universally true. Response dimensions may not improve with practice. Skilled performance typically requires adjusting response speed and amplitude so as to maximize the reinforcement contingent upon that performance. In an experimental situation, Marzocco (8) found that the force exerted by rats on a bar actually decreased with continued training, rather than showing the increase typically found with amplitude measures. And Arnold (1) has found that nonreinforcing bar depressions above an upper limit eliminates most of the stronger depressions.

Secondly, the response dimensions are not always highly correlated. It is well known that similar speeds of performance during acquisition may be followed by markedly different resistance to extinction, depending upon the reinforcement schedule. Furthermore, in Marzocco's experiment, mentioned above, speed *increased* with training while force was *decreasing*.

These implications that the response dimensions should improve with practice and be correlated, and the evidence against them, are well known to most behavior theorists. The solution that Hull apparently chose for this dilemma was to retain the macromolar theory for most behavior, switching to the *micromolar* analysis (which is discussed subsequently) for situations where the former will not work. Of course, this solution can be attempted only if one can identify *in advance* which situations demand use of the micromolar approach. According to Hull (3, p. 339) "It is known that striated-muscle reactions may easily be trained to a particular amplitude by special conditions of reinforcement. . . . . it is assumed in the following analysis that no such complications are involved." That is, use the classical macromolar analysis as long as no response dimension is reinforced per se. In a Skinner box, for example, use the micromolar theory if weaker responses or slower responses, etc., are differentially reinforced; unless such differential reinforcement of some response dimension is involved, use the macromolar system.

Exactly the reverse is a thesis of this paper: the *macromolar* system holds only when the response dimension *is* differentially reinforced, and then *only* under a particular set of conditions of reinforcement. In the Marzocco experiment, force decreased with practice not because medium forces were differentially reinforced with more food, but because increasing forces were not differentially reinforced with more food. A response dimension will improve monotonically with training if reinforcement increases monotonically with that response dimension; the response dimensions will show stable correlations in different situations when the conditions of reinforcement remain the same for those situations. While there are, therefore, many situations for which the macromolar system will work success-

monotonically with practice. A dimension could logically be nonmonotonically related to excitatory potential if it is always so related. The more technically correct statement, therefore, is that all response dimensions must bear invariant relationships to the number of reinforced trials. An analogous technical qualifier should be assumed for the correlation between response dimensions.

fully, the micromolar approach has the advantage of being appropriate for all situations.

The above arguments question the general usefulness of the classical Hullian model within the specific area for which it was designed. Let us turn now to a description of two lines of research [2] which are currently being investigated in the writer's laboratory, and which manipulate variables in a somewhat different manner from that explicit in the macromolar system.

In one of these lines of research, the *conditions-of-reinforcement* project, *S*s are differentially rewarded for different response speeds. That is, more food can be given for slow speeds, or a longer delay can be imposed for fast speeds, or in general any function can be chosen to relate reinforcement to performance. In the other research, the *free-behavior situation, S* lives in a cage, and the amount of water it receives (and presumably therefore its drive) depends upon performance. That is, higher response rates produce more water per unit time.

Response and reward in the conditions-of-reinforcement project, and response and drive in the free-behavior situation, may be viewed as mutually interdependent, whereas the classical Hullian model views response as a dependent variable, with drive and reward among the independent variables. In an attempt to handle problems of the kinds described, two approaches are being explored:

1. A new method of solution within the classical macromolar system may be devised. This will presumably be some form of equilibrium analysis such as, for example, the economic theory based on the mutual interdependence of price and quantity. Thus, when reinforcement varies with speed, the organism will be viewed as running at just that speed which provides just that amount of reward which produces just that degree of incentive which will motivate just that speed. The extent to which this approach will require fundamental changes in the macromolar system is not yet clear. For example, in the free-behavior situation there is no simple identification of what constitutes a "trial," and therefore the meanings of those intervening and dependent variables which involve such a notion are unclear. Further, a straightforward equilibrium solution of the Hullian model appears to imply, for example, that *S*s cannot be trained to perform at different amplitudes for the same amount of reward, which is not consistent with the findings described earlier (1). However, the equilibrium approach is being developed together with the micromolar one.

2. The micromolar level of analysis can be employed. In a micromolar theory, quantitatively different behaviors are viewed as different responses; e.g., running an alley in three seconds is not just a more vigorous macromolar response than a five-second running time, but is a different micromolar response having a separate response tendency. It is the purpose of the next part of this paper to describe this approach.

THE MICROMOLAR APPROACH

The micromolar approach treats quantitatively different levels of response (e.g., different speeds or amplitudes) as different responses. Analysis of the response thus parallels the traditional analysis of the stimulus. Just as different wave lengths and different intensities of a visual stimulus can be discriminated, so different movements and different amplitudes of a bar response are to be differentiated. In a micromolar system, therefore, the re-

[2] Supported by the National Science Foundation. The free behavior situation is being studied in collaboration with B. S. Rosner.

sponse side of the S-R sequence takes on the same general feature as the stimulus side—i.e., that both qualitative and quantitative variations are viewed as different events. Associated with each of these response events is an excitatory tendency from which one can, in principle, calculate the probability of occurrence of any particular one of the events.

It is the aim of this section to discuss a few of the general features of a micromolar system and to illustrate the method of solution. Since a complete micromolar theory cannot be presented at this time, the illustrations will call heavily upon the classical macromolar Hullian model. Lower case subscripts (e.g., ${}_sH_r$) are used rather than capitals (e.g., ${}_sH_R$) to remind the reader that the definition of response has been modified as described above. A number of technical details are given in footnotes; such treatments are admittedly sketchy and are included only for those readers interested in a detailed study of the micromolar approach.

For illustrative purposes, let us consider a rat running down an alley to a food reward. One trial a day is given, with drive, delay of reinforcement, and magnitude of reinforcement held constant. We must first ask why it is that the rat's speed increases when the experimenter apparently is not providing differential reinforcement, e.g., more food for faster speeds.

The micromolar approach requires the addition of new dimensions of reinforcement. One which is suggested is called the *interval* of reinforcement, which is defined as the time between presentation of the stimulus and receipt of the reward.[3] Accordingly, the rat does in point of fact have the interval of reinforcement under his control since faster speeds result in shorter intervals of reinforcement. As the rat runs faster, it gets the food sooner; differential reinforcement is indeed present. The basic assumption, of course, is one of maximization of reward, i.e., the rat will tend to select speeds yielding the greatest reinforcement.

But, the argument goes, the rat must once run faster to experience this shorter interval of reinforcement; and would not the micromolar theory imply that he would continue to make that micromolar response (the slow one) that was first rewarded? The micromolar theorist must here introduce quantitative response generalization. Just as there are quantitative and qualitative stimulus generalizations, so now there are quantitative and qualitative response generalizations.[4] The initially rewarded slow response is the most likely, but faster responses will, by quantitative response generalization, also tend to occur. It is when these faster responses do occur that differential reinforcement is received, making the faster responses relatively prepotent.

We see, then, that the rat occasionally makes faster responses because of response generalization, and that these become more probable than slower responses because they provide shorter intervals of reinforcement. The reader familiar with Hull's writing may recall

[3] I also propose the *latency* of reinforcement as the time between the beginning (or any identifiable stage except the end) of the response and the receipt of the reward. The latency of reinforcement is, of course, the delays of reinforcement of the different parts of the response, but it may be useful to keep conceptually separate the time the S must wait after the response is complete (the delay of reinforcement) from other time periods.

[4] Quantitative response generalization has no direct empirical referent; i.e., one cannot generate a gradient of quantitative response generalization by operations comparable to presenting different stimuli and employing curve-fitting techniques. Quantification of a micromolar system must be done by methods involving successive approximations.

TABLE 1

ILLUSTRATIVE NUMERICAL STEPS IN THE SOLUTION OF A MICROMOLAR THEORY TO PREDICT THE TERMINAL PROBABILITY DISTRIBUTION OF A QUANTITATIVE RESPONSE MEASURE (SPEED)

R'	Interval of Reinforcement	${}_sJ_r$	${}_sK_r$	D	${}_sH_r$	${}_sE_r$	${}_sF_r$	${}_s\bar{E}_r$	p
1	10.00	.10	.67	.90	1.00	.060	.005	.055	.00
2	5.00	.20	.67	.90	1.00	.120	.020	.100	.00
3	3.33	.30	.67	.90	1.00	.180	.045	.135	.01
4	2.50	.40	.67	.90	1.00	.240	.080	.160	.09
5	2.00	.50	.67	.90	1.00	.300	.125	.175	.23
6	1.67	.60	.67	.90	1.00	.360	.180	.180	.34
7	1.42	.70	.67	.90	1.00	.420	.245	.175	.23
8	1.25	.80	.67	.90	1.00	.480	.320	.160	.09
9	1.11	.90	.67	.90	1.00	.540	.405	.135	.01
10	1.00	1.00	.67	.90	1.00	.600	.500	.100	.00

his similar analysis of selection among qualitatively different responses (2).

In order to proceed from this point in a practicable manner, some restriction must be placed on the number of micromolar responses involved in any response dimension. Response speed, for example, may vary over an indefinitely large number of values, and if one had to calculate an ${}_sE_r$ for each of these for each trial, the labor involved would be prohibitive. I propose, therefore, that we initially try out a system which has only 10 micromolar levels of each response dimension.[5]

Let the symbol R' stand for micromolar response, of which there are, by convention, 10 in the speed dimension. These are listed in the first column of Table 1. We wish now to calculate the terminal excitatory potential for each of these R's in order to illustrate how the rat's speed is predicted by a micromolar model. This illustration will involve certain arbitrary assumptions concerning the quantitative characteristics of the functions. It would be fortuitous if any of these are correct, and the reader should not assume, for example, that a symmetrical probability distribution is a *necessary* consequent of this

[5] The fact that speed, as recorded, varies continuously does not imply that indefinitely small gradations are psychologically different responses, any more than a continuous physical measure of stimulus intensity fits the psychological perceptions. While an analogue to the stimulus jnd may be devised which may suggest that there are more than 10 micromolar levels (depending upon the response system and the dimension involved), this first approximation will suffice for the present. Even so, if one admits to only four response dimensions, each of which has 10 micromolar levels, there are 10,000 different micromolar responses. This can be simplified by assuming that the response dimensions are independent, such that separate calculations can be made for each dimension. That is, the system can be solved for amplitude where that is being recorded, for speed where that is being recorded, or for them both *separately*, when both are being recorded.

This assumption of independence among the response dimensions, it will be recognized, is in sharp contrast with the macromolar assumption of perfect correlation between them. The actual facts may well turn out to be intermediate; i.e., high-amplitude responses evoked by differential reinforcement may encourage speeds faster than expected from the speed conditions of reinforcement alone. Because experimenters typically record only one response dimension, and because we do not yet know the possible interdependence among the response dimensions, I would propose the independence assumption as the most fruitful starting point. Statements of interdependence can subsequently be added when data are collected which require them.

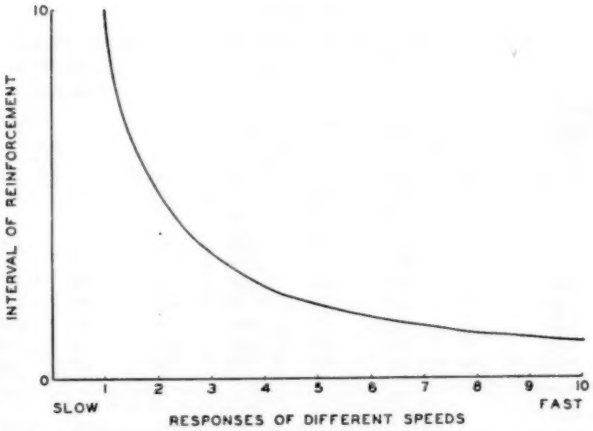

FIG. 1. Function showing how the interval of reinforcement (the time between presentation of the stimulus and receipt of the reward) is related to responses of different speeds (different micromolar responses); i.e., *S* gets the food sooner the faster it runs.

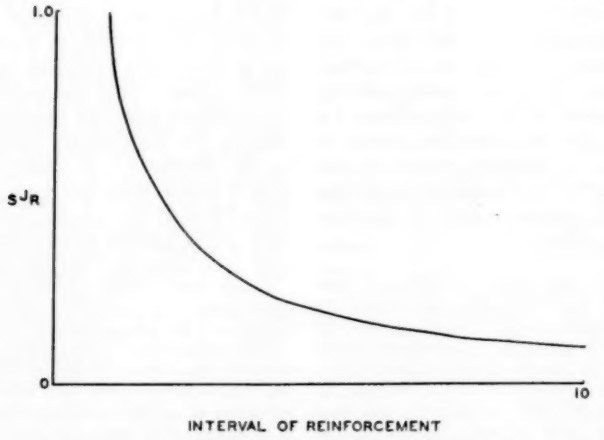

FIG. 2. Function showing how the incentive variable ($_{s}J_{r}$) is determined by the interval of reinforcement. This function was arbitrarily selected for purposes of illustration.

micromolar approach. The illustration is intended to justify further development of the model to the point where it can be tested more rigorously against data from a program of quantitative research.

Assuming that the incentive variable $_{s}J_{r}$ [6] is based on the interval of rein-

[6] The symbols $_{s}J_{r}$ and $_{s}K_{r}$ are used for Hull's *J* and *K* in order to imply their specificity to a particular *R*′. While, following Spence (11), these intervening variables are

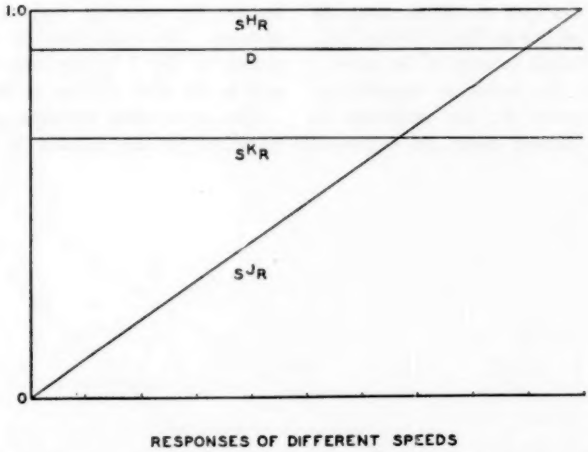

FIG. 3. Functions showing the terminal values of the intervening variables for each micromolar response. Habit ($_sH_r$), drive (D), and incentive ($_sK_r$) are assumed in this case to be independent of speed. The other incentive variable ($_sJ_r$) is determined by multiplication of the functions in Figs. 1 and 2.

forcement, Fig. 1 shows the condition of reinforcement relating speed to interval of reinforcement. Appropriate values of this function are given in the second column of the table. This, as suggested earlier, shows that faster running responses (R') produce shorter intervals of reinforcement.

In order to calculate $_sJ_r$ for each R', we must know how $_sJ_r$ is determined by the interval of reinforcement. For illustrative purposes we will take the function in Fig. 2. Now Figs. 1 and 2 yield the linear increasing function labeled $_sJ_r$ in Fig. 3. That is, one determines from Fig. 1 what the interval of reinforcement will be for some particular R', and then determines from Fig. 2 what value of $_sJ_r$ is produced by that interval of reinforcement. Successive solution for each R' generates the functional relation between $_sJ_r$ and R', as shown in Fig. 3, and listed in the third column of Table 1.

The other incentive variable, $_sK_r$, is based upon the magnitude of reinforcement. In the type of design described, magnitude of reinforcement is constant no matter what the S's speed; i.e., speed is not differentially rewarded with different magnitudes of reinforcement. Accordingly, the $_sK_r$ variable will be the same for all R's, since each receives the same magnitude of reinforcement. For illustrative purposes, we shall take $_sK_r$

acquired, presumably on the basis of the fractional anticipatory goal response, it will be simplest at this stage to use the initial Hullian approach of relating incentive directly to reinforcement without this mediating link. This method has the unintended implication that incentive is not gradually acquired, but it will nonetheless permit satisfactory discussion of limiting values.

One further disadvantage arises from failure to specify the learned characteristic of incentive. This is because, in a micromolar theory, the incentive will not exactly parallel the conditions of reinforcement, since the incentive variable will itself generalize to similar R's, and the effect will be to smooth out, in the incentive function, discontinuities that may prevail in the conditions of reinforcement.

at an arbitrary level of .67, as shown by the horizontal line in Fig. 3 labeled ${}_sK_r$, and in the fourth column of the table.

Similarly, the hours of deprivation, which determine D, are unrelated to R'; i.e., S cannot affect its degree of hunger by selecting fast or slow responses. We shall take D as .90, as shown in Fig. 3 by the line labeled D, and in the fifth column of the table.

The associative variable, ${}_sH_r$, is determined by the number of reinforced

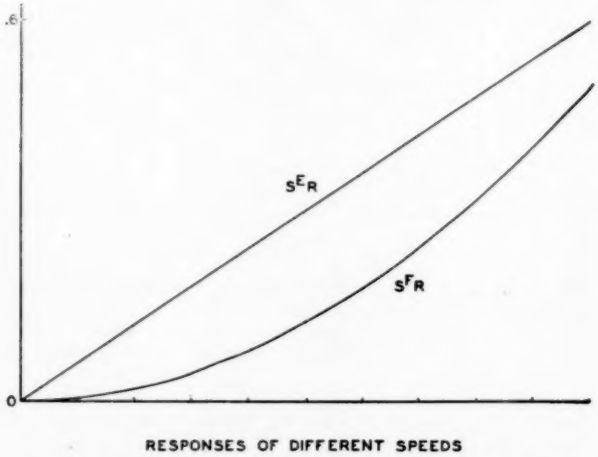

FIG. 4. Excitatory tendency (${}_sE_r$) and effort (${}_sF_r$) for each micromolar response. ${}_sE_r$ is the product of the four functions in Fig. 3. The ${}_sF_r$ function was arbitrarily selected for purposes of illustration.

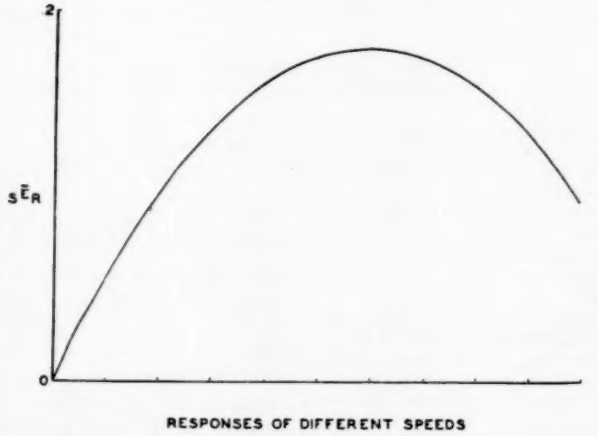

FIG. 5. Effective excitatory tendency (${}_s\bar{E}_r$) for each micromolar response, obtained as the subtraction of ${}_sF_r$ from ${}_sE_r$ in Fig. 4.

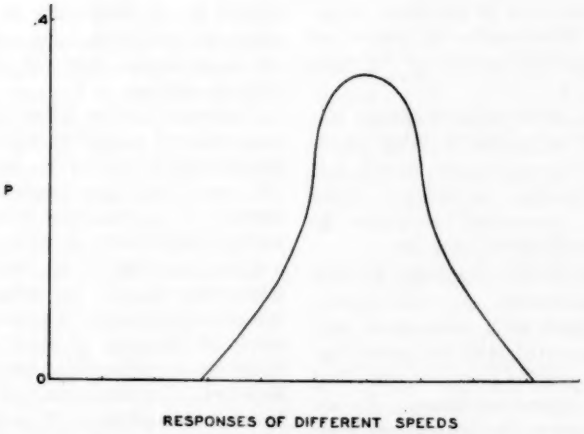

FIG. 6. Terminal distribution showing the probability (p) that any particular micromolar response will occur. This was obtained by applying oscillation ($_sO_r$) to $_s\bar{E}_r$.

trials at that particular R' or any similar R' from which $_sH_r$ generalizes. $_sH_r$ will have changed during training in a pattern not uniquely specifiable, but we may assume that enough practice will have occurred for the $_sH_r$ of each R' to have attained the limit of unity. This is recorded in the sixth column of the table, and an appropriately labeled function is shown in Fig. 3.

Excitatory tendency ($_sE_r$) is assumed to be a multiplicative function of the four variables $_sJ_r$, $_sK_r$, D, and $_sH_r$. To calculate $_sE_r$ for each micromolar response, therefore, we need only multiply the four functions in Fig. 3. This operation yields the values in the seventh column of the table. These are graphed as the linear function labeled $_sE_r$ in Fig. 4.

At this stage a new variable must be introduced. Faster speeds are assumed to require more effort which should act to reduce the likelihood of selecting fast responses, other things equal. Let us use the symbol $_sF_r$ for this effort variable,[7] and assume that it increases with increased speed according to the function labeled $_sF_r$ in Fig. 4. Appropriate values from this function are recited in the eighth column of the table.

Effective excitatory tendency ($_s\bar{E}_r$) is calculated as the subtraction of $_sF_r$ from $_sE_r$. This was done, and the result is shown in the ninth column of the table and in Fig. 5. We have therefore arrived at the $_s\bar{E}_r$'s for each of the R''s, to each of which behavioral oscillation must yet be applied. We therefore make one last transformation, convert-

[7] Work, in the macromolar Hullian analysis, produces reactive inhibition which competes with $_sE_r$ after the work has been done and before the inhibition is dissipated. While reactive inhibition should be retained, the micromolar approach requires that the S also be conceived as anticipating the work involved in each micromolar response. This $_sF_r$ variable suggested will probably be successfully related to a fractional anticipatory resting response, just as $_sJ_r$ and $_sK_r$ are related to the fractional anticipatory goal response (because $_sF_r$ must also be a learned variable), but we will here content ourselves with its simple postulation.

ing ${}_{s}\bar{E}_{r}$ values into probabilities of occurrence. The results of this[8] are shown in the last column of the table and in Fig. 6.

In Fig. 6 we therefore have the distribution of asymptotic running speeds expected by the micromolar system, and one may calculate the weighted mean of these for a convenient prediction. In this case, predicted speed is .60.

Let us extend this illustrative solution just one step further. At casual glance, it would appear as if phenomena such as the gradient of stimulus generalization would provide a major stumbling block to a micromolar theory. To see that this is not so, the reader may refer again to Fig. 3, and visualize a shift in the ${}_{s}H_{r}$ function down from unity to some lower value due to a generalization decrement. Simply tracing the results of this will show that the ${}_{s}E_{r}$ function in Fig. 4 would shift down to a line of lesser slope, the ${}_{s}E_{r}$ function in Fig. 5 will shift downward and to the left, and the point of maximum probability in Fig. 6 will shift toward a slower speed R'. If this is done numerically, a change of ${}_{s}H_{r}$ from unity to .50 would result in a weighted mean speed of .30, which is slower than the .60 of the original solution.

Analogous results follow changes in magnitude of reinforcement or hours of deprivation, which lead to shifts in the ${}_{s}K_{r}$ or D functions, respectively, and thereby to appropriate shifts in the terminal probability distribution.

Space does not permit describing the micromolar solution for other types of behavior situations. However, an example of the type of novel prediction made by a micromolar theory can be sketched. Consider a rat that is trained to run in two alleys, at the end of which it receives one food pellet. To receive the pellet in Alley A, its speed must exceed an arbitrary level; in Alley B, reward follows only if its speed is less than this level. Under such conditions of reinforcement, if the two alleys are discriminably different, the micromolar approach implies that the rat will learn to run relatively fast in Alley A and relatively slow in Alley B. Further, if, following such training, with all other things constant, the rat is given a choice between Alleys A and B, the prediction is that it will tend to choose Alley B (the one in which it runs slowly). This follows because Alley A responses will involve greater effort, and the distribution of their effective excitatory potentials will be lower relative to that of the Alley B responses. Since the experiment has not as yet been performed, it represents a prediction by this micromolar model.

It can be noted that the general macromolar assumption that choice behavior is related to speed implies that rats will choose the alley in which they run the faster. This is the opposite of the micromolar prediction under the given conditions.

[8] I do not propose an exact mathematical solution; the solution used here preserves the order of relationships which must also result from an exact solution. Probabilities were first determined for each pair of R''s, and assembled in a matrix showing the probabilities that any R' will exceed any other R', i.e., the probability that R_1 exceeds R_2, that R_1 exceeds R_3, etc. The probability that R_1 exceeds R_2, R_3, . . . and R_{10}, is provided by the multiplication of row entries on the assumption of independence. This assumption is, of course, not correct in this situation, and the calculated probabilities will underestimate the true probability such that the total probability of any R' will be less than unity. This residual was simply distributed among the R's in proportion to the probability attained on the assumption of independence. A more exact solution will not be required until the system has reached the maturity of quantification.

SUMMARY

1. A discussion of the fundamental features of the classical macromolar approach to behavior theory, as exemplified by Hull, suggested (*a*) that the implications that all response dimensions should improve with practice and be highly correlated are not universally true; and (*b*) that a new method of solution of the model is necessary for situations in which reward or drive depends differentially upon some response dimension, such as speed or amplitude.

2. A micromolar approach was described, in which quantitatively different behaviors (e.g., different amplitudes) are viewed as different responses, rather than as different strengths of the same response.

3. An illustrative micromolar model was solved for a simple instrumental response, namely, a rat's speed of running down a straight alley under several conditions of reinforcement. A prediction was made by this model concerning choice between two differentially rewarded responses.

4. It is concluded that the micromolar approach is a potentially useful one, and deserves consideration alongside other theories of learned behavior.

REFERENCES

1. ARNOLD, W. J. An exploratory investigation of primary response generalization. *J. comp. Psychol.*, 1945, **38**, 87–102.
2. HULL, C. L. The concept of the habit-family hierarchy and maze learning. *Psychol. Rev.*, 1934, **41**, 33–52 and 134–152.
3. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
4. HULL, C. L. *Essentials of behavior*. New Haven: Yale Univer. Press, 1951.
5. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
6. LOGAN, F. A., *et al. Behavior theory and social science*. New Haven: Yale Univer. Press, 1955.
7. LOGAN, F. A., BEIER, EILEEN M., & ELLIS, R. A. The effect of varied reinforcement on speed of locomotion. *J. exp. Psychol.*, 1955, **49**, 260–266.
8. MARZOCCO, F. N. Frustration effect as a function of drive level, habit strength and distribution of trials during extinction. Unpublished doctor's dissertation, State Univer. of Iowa, 1951.
9. SIDMAN, M. The temporal distribution of avoidance responses. *J. comp. physiol. Psychol.*, 1954, **47**, 399–402.
10. SKINNER, B. F. *The behavior of organisms*. New York: D. Appleton-Century, 1938.
11. SPENCE, K. W. Theoretical interpretations of learning. In C. P. Stone (Ed.), *Comparative psychology*. New York: Prentice-Hall, 1951.

(Received March 21, 1955)

Psychological Review
Vol. 63, No. 1, 1956

SUBCEPTION: FACT OR ARTIFACT? [1]

CHARLES W. ERIKSEN [2]

Harvard University

Several years ago Lazarus and McCleary (6) reported an experiment in which GSR's were conditioned to one group of nonsense syllables and not to another group. They found that a tachistoscopic exposure of a conditioned nonsense syllable resulted in a larger GSR than presentation of a nonconditioned syllable, even when the duration of the exposure was too rapid for correct verbal identification. Lazarus and McCleary termed this effect "subception," which they defined as "a process by which some kind of discrimination is made when the subject is unable to make a conscious discrimination." These authors do not clarify the nature of this process, but their definition and their interpretation of their data imply that the autonomic nervous system, under certain circumstances at least, is capable of making more accurate discriminations among stimuli than exist in the individual's awareness.

This interpretation of their data has received criticism, and alternative interpretations have been offered. Bricker and Chapanis (1) have shown that, following an incorrect verbal response to a tachistoscopically exposed stimulus, the mean number of additional guesses required to name the stimulus correctly was significantly less than would be expected on the basis of random guessing. They argue from this finding that, even when the stimulus was incorrectly responded to, subjects obtained fragmentary cues that limited the number of plausible interpretations or responses. If a conditioned nonsense syllable is among the plausible responses, they expect the GSR to be greater than if the plausible responses contained only nonconditioned syllables. Since the exposure of a conditioned syllable would always or nearly always result in this syllable being among the plausible responses, GSR's would then be expected to average larger on such trials than on trials where a nonconditioned syllable was presented.

Howes (3) has also presented an account of the subception effect based upon a statistical model of discrimination. His alternative to a subception process is the hypothesis that "at any specified moment the GSR accompanying an observer's report is proportional to the probability that that report will be a shock [conditioned] syllable" (p. 99).

Both Howes (3) and Bricker and Chapanis (1) have dealt with important aspects of the subception phenomena, but they have not succeeded in coming to grips with certain logical fallacies and operational limitations of the subception experiment. In fact, the present author has heard it argued on several occasions that the Bricker and Chapanis experiment actually confirms the subception process: "After all they demonstrated unconscious discrimination, since they showed that even when the subject was unable to report the correct stimulus, if you forced him to 'guess' he was right more often than would be expected by chance." Even Howes does not attempt to refute the subception interpretation, but presents his statistical model as an alternative

[1] This paper is the outgrowth of a research program supported by a grant from the Laboratory of Social Relations, Harvard University.

[2] Now at The Johns Hopkins University.

description (a model, incidentally, which one could use as a sophisticated account of unconscious discrimination, since Howes does not provide the operations for determining whether the distribution of response probabilities in a subject is represented in awareness or not).

It seems to the present author that much of the confusion and controversy over subception and problems in "awareness" in general is due to unclear formulations of the questions, and to failure to consider the converging operations (2) necessary to permit interpretations of data in terms of unconscious discrimination. It is the purpose of the present paper to reformulate and reanalyze the subception experiment in terms of partial correlation between the stimuli and the two responses, and to show that, when so conceived, the Lazarus and McCleary results are not only consistent with but would be predicted by well-known experimentally derived facts. It is a further purpose of the paper to make clear that the invoking of a subception process to explain the Lazarus and McCleary results was not only unnecessary but was unjustified from the experimental operations employed.

The subception effect as a case of partial correlation. The results of the Lazarus and McCleary experiment can be stated in terms of partial correlation.[3] The essence of their experimental design was the requirement that the subject make two concurrent discriminated responses to the perceptual stimulus, namely, a GSR and a verbal response. They found that when the verbal response was held constant, the partial correlation between the stimulus and the GSR was significantly greater than zero. This partial correlation was termed the subception effect by the experimenters, and its occurrence was attributed to the operation of unconscious discrimination.[4]

However, before we can accept this interpretation, we need to examine the general circumstances under which such a partial correlation could occur. We can define the general case of what Lazarus and McCleary have termed the subception effect as the case where the partial correlation between the stimulus and one response, with the second response held constant, is some positive value greater than zero and less than one. The general circumstances under which such a partial correlation would be obtained are where the two responses, R_1 and R_2, are each correlated with the stimulus, but less than perfectly and less than perfectly correlated with each other. In other words, we have a case where each response or response system has a finite error term associated with it, and these error terms are less than perfectly correlated between the two response systems. These conditions are not only sufficient but actually determine that a positive partial correlation will be obtained between the stimulus and either response with the remaining response held constant.

It should be noted that these are essentially the conditions that obtained in the Lazarus and McCleary experiment. They arranged the tachistoscopic duration of stimulus exposures so that the verbal report of the subjects was only about 50 per cent accurate, or so that the correlation of the verbal response with the stimulus was considerably less than perfect. The GSR in their experiment was also less than perfectly correlated

[3] The term correlation as used in this paper does not refer to any particular correlation statistic. For this discussion, however, it will be assumed that all correlations are positive or nonmetric.

[4] The distinction between the subception effect and the subception process previously made by Howes (3) will be observed in this paper.

with the stimulus. From previous research we know that the GSR is an unstable response, and the Lazarus and McCleary data also show that large GSRs were obtained to nonconditioned syllables. Since these investigators actually obtained a significant partial correlation between the stimulus and the GSR, we know that the remaining condition (less than a perfect correlation between the two responses) must have existed.

When we formulate the results of the Lazarus and McCleary experiment in partial correlation terms, it becomes possible to state clearly and concisely what their experiment demonstrated. They succeeded in showing that both the verbal response and the GSR are each, at least to a degree, independently correlated with the stimulus and are partially independent of each other. In other words, the GSR is at least partly determined by the stimulus and the intervening perceptual process, and not solely by the verbal response.

The nature of the Lazarus and McCleary result can be seen more clearly if we diagram two possible relations between the verbal response and the GSR.

A. Stimulus → Perceptual Process → Verbal Response → GSR

B. Stimulus → Perceptual Process → Verbal Response / GSR

In A we have the condition where the GSR is dependent upon the verbal response to the perception, while in B we have two parallel response systems, each separately dependent upon the perceptual processes. If the relation between verbal response and GSR was as diagrammed in A, it would not have been possible for Lazarus and McCleary to have obtained their results. Such a relation between verbal response and GSR would permit less than a perfect correlation between them owing to error that might exist in the GSR, but it would have been impossible to have obtained a true correlation between the GSR and the stimulus with the verbal response held constant. The fact that these investigators found a partial correlation between the stimulus and GSR when the verbal response was held constant demands a relationship similar to that diagrammed in B, unless their results are to be attributed to sampling fluctuation.

Some factors contributing to independence between GSR's and verbal responses. It should be carefully noted that the demonstration that both the verbal response and the GSR are independently related to the stimulus is not the same as demonstrating unconscious perception or discrimination. All we have evidence for so far is the presence of two parallel response systems, neither of which is perfectly reliable, and whose error terms are not perfectly correlated.

In view of the neurophysiological differences between verbal responses and GSR's, it is not surprising that the two response systems are not perfectly correlated. Not only are there differences in the latencies for the central nervous system and autonomic nervous system responses, but the motor systems also differ. Any differences in the neurophysiology of the responses allow, theoretically at least, certain physiological variables to affect one response system without affecting the other.

The existence and effect of such physiological variables upon responses were formally recognized by Hull (5) in his principle of behavior oscillation. This principle acknowledged certain random variations in responses and response strength that Hull felt might be attributed to transitory variation in physiological states. He has described the empirical evidence for this principle

(5), along with evidence indicating that this type of response error was apt to be uncorrelated between different responses; he referred to the latter as *asynchrony of oscillation*. The fact that Hull found it necessary formally to recognize noncorrelated errors between different responses indicates that their presence in the subception effect is not a novel event.

While the principle of behavior oscillation attests to the empirical basis for assuming noncorrelated error terms between the verbal and autonomic response systems, it does not specify variables that might differentially affect the two systems. However, certain differences in the nature of verbal and autonomic responses suggest some psychological variables that might be expected to produce errors in verbal responses independently of the GSR.

The GSR is essentially a single response varying only in magnitude, and is analogous to a verbal response system with only a single verbal response that can vary in response strength. If the verbal response system contains two or more responses, this permits new variables that may affect the verbal discrimination while leaving the GSR unaffected.

Let us take as an example an experiment in which we have two nonsense syllables, one of which has been conditioned as a stimulus for a GSR. In addition, our subject in this experiment has two verbal responses, one for each nonsense syllable. If we now expose the nonsense syllables tachistoscopically, at durations too short for errorless discrimination, we could reasonably expect some disagreement in the errors made in the response systems. The two verbal responses might not be of the same response strength, so that one of them might be evoked more often than the other. Also, on some trials on which the correct response would have the greatest habit strength to the exposed stimulus, the incorrect response with less habit strength to that stimulus might actually be evoked, due to response competition and behavior oscillation. Response variables such as differing habit strengths and competition could contribute to errors in the verbal responses while leaving the GSR unaffected.

Effect of the number of response categories upon the subception effect. If the number of response categories in one response system is limited relative to the other response system, the correlation between the two response systems will be reduced, as will also the correlation of the restricted response system with the stimulus. This situation will also contribute to a large partial correlation between the stimulus and the unrestricted response system. In comparing verbal responses with GSR's we are comparing a response system with a discrete distribution with one that for all practical purposes is continuously distributed. If the number of verbal response categories is too few, we run the risk of spuriously reducing the correlation between the verbal response and the stimuli, and of increasing the partial correlation between the stimuli and the GSR.

As an example, let us suppose an experiment in which we have 10 different-sized squares, to the largest of which we have conditioned a GSR. If the squares are now presented to the subjects one at a time, in accord with stimulus generalization (4, 8) we would expect the magnitude of GSR to vary with the square size, the largest magnitude occurring to the largest square and the smallest magnitude to the smallest square. In addition to recording the GSR on each stimulus presentation, let us also require the subject to make a verbal judgment of the square size, for which we allow him two response categories,

"large" and "small." If our subject called the five largest squares "large," we would have no indication from his verbal behavior that he could discriminate among these five squares. Yet we might reasonably expect that the magnitude of GSR would vary with square size within this single verbal response category. In this case we would have a large partial correlation between the stimuli and the GSR when the verbal response was held constant.

Few would be willing to accept the results of our hypothetical experiment as demonstrating unconscious discrimination, without the control operation of permitting the subject to use at least 10 verbal responses. Yet a similar situation prevailed in the Lazarus and McCleary experiment, which contributed an unknown amount to the subception effect they obtained. In their experiment subjects were allowed to use only one of 10 nonsense syllables as a response. While 10 response categories may seem like an adequate number, their adequacy depends upon the complexity of the perceptual stimulation the subjects are to receive.

While we tend to act as though the perceptual stimulus presented to a subject in a tachistoscopic exposure of a nonsense syllable was the nonsense syllable itself, this is not the stimulus perceived by the subject. If it were, there would be no point in using the tachistoscope. When the nonsense syllable SIDAK is tachistoscopically exposed for a short duration and the subject is instructed to report just what he perceives (and is allowed the freedom of the English language for responses), we tend to get responses such as "I saw an 's' followed by what seemed to be two blurred letters, then an 'a' and a 'k'." If, however, the subject is instructed implicitly or explicitly to respond in terms of a preselected list of nonsense syllables, we obtain the response "sidak" or some other nonsense syllable from the permitted list that has the correct sequence of letters. A second exposure of this nonsense syllable might result in the subject's perceiving all of the letters. If the subject's responses were restricted to preselected nonsense syllables, both exposures might lead to the same verbal response of "sidak," whereas with greater response freedom we would be able to distinguish two different perceptual experiences.

The manner in which verbal response restriction could have contributed to the subception effect obtained by Lazarus and McCleary can be illustrated as follows. Assume that a subject has been exposed to the nonsense syllables VAVUK, YUVUF, BEKUM, and LUFED, and that a GSR has been conditioned to only VAVUK. If now the conditioned syllable VAVUK is tachistoscopically exposed for a short duration, the subject under conditions of free report might respond by saying that he saw only a dark blur with an impression of a "u" somewhere in the blur. A second exposure of the conditioned syllable might produce the report that it looked like a blur again, but that the next to the last letter was a "u." To a third exposure the response might be that he saw blurred letters with a "vu" in the center. If the subject in our example was constrained to respond with one of the four nonsense syllables, all three exposures might well have led to the response "yuvuf."

There is reason to believe that stimulus generalization of the GSR would occur to these fragmentary tachistoscopic perceptions (7). If so, the magnitude of the GSR would increase progressively through these three exposures as the subject's perception, as revealed by free report, became more similar to the conditioned syllable. In this case the magnitude of the GSR would correspond quite closely with the free verbal re-

port. However, if the subject was permitted to use only nonsense syllables as responses, he might group these three different perceptions under a single response. Thus, we have the same condition that existed in our hypothetical experiment on square sizes. We have forced the subject to group what may be differentially reportable perceptions under a common verbal response. We have experimentally constrained or reduced the precision of the verbal response system while leaving the GSR unconstrained; a condition which will increase the partial correlation between the stimuli and the GSR while artifactually reducing the correlation of the stimuli with the verbal responses.

Conscious awareness as a discriminated verbal response. It is recognized that there is nothing incompatible between the formulation of the subception effect as a partial correlation and an interpretation of unconscious discrimination. After all, such a partial correlation may be what we have in mind at times when we refer to unconscious discrimination. Variables such as behavior oscillation and response competition which produce noncorrelated errors between two response systems may be among the reasons that all discrimination is not conscious. But it should be clear from the previous section that if one wishes to take this position, it is necessary that he demonstrate that the observed partial correlation is not solely due to artificial restrictions upon the verbal discriminatory system.

Implicit in the Lazarus and McCleary definition of subception is the equation of conscious awareness with a discriminated verbal report. Such a definition of awareness requires an operational control of any experiment attempting to demonstrate unconscious discrimination. This control consists of showing that the experiment provided the subjects with enough different verbal responses to identify separately each perceptual discrimination they were capable of making. In our hypothetical experiment on square sizes, it would not be justifiable to conclude that our subjects were only capable of making two verbal discriminations of size in the range of sizes tested. Before such a conclusion could even begin to be justified, we would need to run a control condition where subjects were permitted to use 10 or more different verbal responses.

To accept Lazarus and McCleary's conclusion that discrimination was shown by the GSR in the absence of conscious awareness, we must be prepared to assume that a language consisting solely of 10 nonsense syllables is sufficient to describe the entire realm of subjective awareness. In fact, however, the tenability of such an assumption has been disproven by Bricker and Chapanis (1). In addition to recording the subject's first response to a tachistoscopic exposure of a nonsense syllable, they also required the subject to give several alternative guesses. These guessing sequences constitute an increase in the number of verbal responses permitted the subjects. When these guessing sequences were taken into account, it was found that verbal discrimination was significantly greater than was indicated when only the subject's initial response was used. In other words, if only the subject's initial response was considered, the amount of conscious discrimination (as defined by Lazarus and McCleary) would be significantly underestimated.

SUMMARY

The present paper has shown that the subception effect reported by Lazarus and McCleary can be formulated as a partial correlation between the GSR and the stimulus with the verbal response held constant. The general case where

such a partial correlation would be obtained is where both the GSR and verbal response are independently correlated with the stimulus, but not perfectly, and the error terms in these two response systems are not perfectly correlated with each other. Several sources of noncorrelated error between the GSR and the verbal response systems were considered.

It was argued that the Lazarus and McCleary experiment did not contain the necessary operations and controls to permit an interpretation of their data in terms of discrimination without verbal awareness. Such an interpretation requires the demonstration that the number of verbal responses available to the subject was sufficient to reflect all the discrimination he was capable of making. The 10 nonsense syllables permitted the subjects as responses in the Lazarus and McCleary experiment were not sufficient to determine the subjects' verbal discrimination capacity.

REFERENCES

1. BRICKER, P. D., & CHAPANIS, A. Do incorrectly perceived tachistoscopic stimuli convey some information? *Psychol. Rev.*, 1953, **60**, 181–188.
2. GARNER, W. R. Context effects and the validity of loudness scales. *J. exp. Psychol.*, 1954, **48**, 218–224.
3. HOWES, D. A statistical theory of subception. *Psychol. Rev.*, 1954, **61**, 98–110.
4. HULL, C. L. The problem of primary stimulus generalization. *Psychol. Rev.*, 1947, **54**, 120–134.
5. HULL, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
6. LAZARUS, R. S., & MCCLEARY, R. A. Autonomic discrimination without awareness: a study of subception. *Psychol. Rev.*, 1951, **58**, 113–122.
7. PAVLOV, I. P. *Conditioned reflexes.* (Trans. by G. V. Anrep) London: Oxford Univer. Press, 1927.

(Received for early publication June 28, 1955)

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by JAMES DREVER

Vol. XLVI. Part 3 August 1955 $3.50 net.

B. A. FARRELL. On the limits of experimental psychology.

JOHN COHEN and C. E. M. HANSEL. The idea of independence.

PER SAUGSTAD. Problem-solving as dependent on availability of functions.

K. LOVELL. A study of the problem of intellectual deterioration in adolescents and young adults.

A. W. HEIM. Adaptation to level of difficulty in intelligence testing.

JOSEPH SANDLER. A test of the significance of the difference between the means of correlated measures, based on a simplification of Student's *t*.

CORRESPONDENCE.

EDITORIAL NOTE.

PUBLICATIONS RECENTLY RECEIVED.

Vol. XLVI. Part 4 November 1955 $3.50 net

OBITUARY NOTICE: Sir Godfrey Thomson. 1881–1955

R. B. CATTELL and J. E. DREVDAHL. A comparison of the personality profile (16 P.F.) of eminent researchers with that of eminent teachers and administrators, and of the general population.

R. G. HOPKINSON. Subjective judgments—some experiments employing experienced and inexperienced observers.

HANNAH STEINBERG. Changes in time perception induced by anaesthetic drug.

S. LANER. Some factors influencing the effectiveness of an instructional film.

T. W. GERVAIS. Freud and the culture-psychologists.

A. B. FITT. A study of racial attitudes during and after the war by the Thurstone technique.

D. B. BROMLEY. Notes on the Shaw Test.

PUBLICATIONS RECENTLY RECEIVED.

OTHER PUBLICATIONS RECEIVED.

The subscription price per volume, payable in advance, is $10 net (post free).

Subscriptions may be sent to any bookseller or to the

CAMBRIDGE UNIVERSITY PRESS

Bentley House, Euston Road, London, N. W. 1